

An Introduction to the Proximal Galerkin Method¹

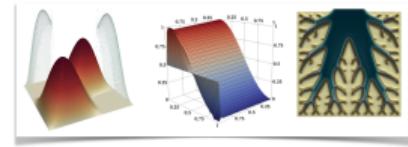
Thomas M. Surowiec, Brendan Keith (Brown U)

Department of Numerical Analysis and Scientific Computing
Simula Research Laboratory
Oslo, Norway

SCAN Meeting, December 7, 2023

¹ Recently mentioned in Popular Science(!): *These 10 scientists are on the cusp of changing the world, It's the Brilliant 10 class of 2023.*
<https://www.popsci.com/science/brilliant-10-2023/>

Overview



1. History and Background
2. Algorithms and Numerical Experiments
3. Extensions of Proximal Galerkin
4. The Latent Variable Proximal Point Method

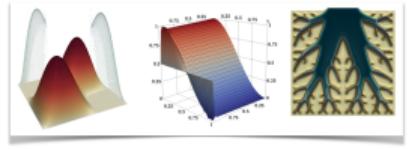


B. KEITH AND T.M. SUROWIEC.

Proximal galerkin: A structure-preserving finite element method for pointwise bound constraints.

Submitted (2023), <https://arxiv.org/pdf/2307.12444.pdf>

Dirichlet's Principle²



In contemporary language, this energy principle states that for all functions $f \in L^2(\Omega)$ and $g \in H^1(\Omega)$, the (weak) solution of Poisson's equation over a Lipschitz domain $\Omega \subset \mathbb{R}^n$,

$$-\Delta u = f \quad \text{in } \Omega, \quad u = g \quad \text{on } \partial\Omega, \tag{1}$$

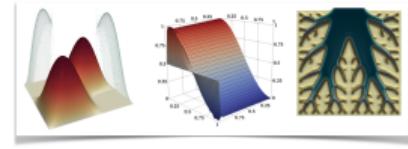
can be obtained as the $H^1(\Omega)$ -minimizer of the Dirichlet energy,

$$E(v) = \frac{1}{2} \int_{\Omega} |\nabla v|^2 dx - \int_{\Omega} vf dx, \tag{2}$$

confined to the constraint set $H_g^1(\Omega) = g + H_0^1(\Omega) = \{v \in H^1(\Omega) \mid v = g \text{ on } \partial\Omega\}$.

²Actually discovered by William Thomson, 1st Lord Kelvin 1847 and C.F. Gauß. Named after his teacher, P.G.L. Dirichlet, by G.F. Riemann in 1900.

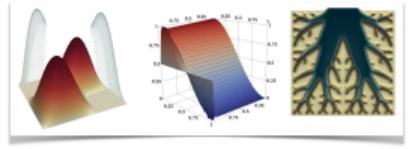
A Variational Inequality



Since $H_g^1(\Omega)$ is nonempty, closed, and convex, the Lions–Stampacchia theorem (1967) states that the energy minimizer $u^* \in K = H_g^1(\Omega)$ is the unique solution to the variational inequality (VI)

$$\int_{\Omega} \nabla u^* \cdot \nabla (v - u^*) dx \geq \int_{\Omega} f(v - u^*) dx \text{ for all } v \in K.$$

A Variational Inequality

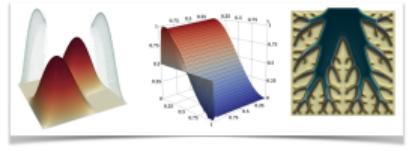


Since $H_g^1(\Omega)$ is nonempty, closed, and convex, the Lions–Stampacchia theorem (1967) states that the energy minimizer $u^* \in K = H_g^1(\Omega)$ is the unique solution to the variational inequality (VI)

$$\int_{\Omega} \nabla u^* \cdot \nabla (v - u^*) dx \geq \int_{\Omega} f(v - u^*) dx \text{ for all } v \in K.$$

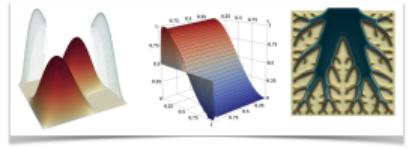
But no one solves Poisson's problem (1) as a variational inequality. Why not?

Back to the Dirichlet's Principle



- $K = H_g^1(\Omega)$ is affine and for all $w \in H_0^1(\Omega)$ and $v \in H_g^1(\Omega)$, $v + w \in H_g^1(\Omega)$, as well.

Back to the Dirichlet's Principle



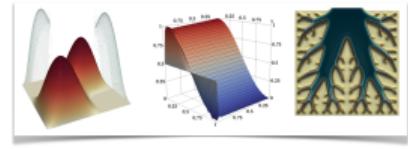
- $K = H_g^1(\Omega)$ is affine and for all $w \in H_0^1(\Omega)$ and $v \in H_g^1(\Omega)$, $v + w \in H_g^1(\Omega)$, as well.
- Taking $v = u^* \pm w$ for any $w \in H_0^1(\Omega)$ here

$$\int_{\Omega} \nabla u^* \cdot \nabla (v - u^*) dx \geq \int_{\Omega} f(v - u^*) dx \text{ for all } v \in K,$$

brings us back to Dirichlet's principle: Find $u^* \in H_g^1(\Omega)$ such that

$$\int_{\Omega} \nabla u^* \cdot \nabla w dx = \int_{\Omega} f w dx \text{ for all } w \in H_0^1(\Omega).$$

Back to the Dirichlet's Principle



- $K = H_g^1(\Omega)$ is affine and for all $w \in H_0^1(\Omega)$ and $v \in H_g^1(\Omega)$, $v + w \in H_g^1(\Omega)$, as well.
- Taking $v = u^* \pm w$ for any $w \in H_0^1(\Omega)$ here

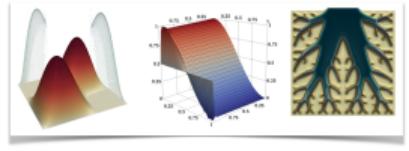
$$\int_{\Omega} \nabla u^* \cdot \nabla (v - u^*) dx \geq \int_{\Omega} f(v - u^*) dx \text{ for all } v \in K,$$

brings us back to Dirichlet's principle: Find $u^* \in H_g^1(\Omega)$ such that

$$\int_{\Omega} \nabla u^* \cdot \nabla w dx = \int_{\Omega} f w dx \text{ for all } w \in H_0^1(\Omega).$$

We use the explicit *geometry* of the feasible set to derive a simpler problem.

The Obstacle Problem³



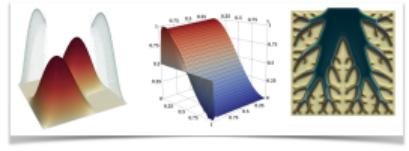
- Now minimize Dirichlet's energy over the set K defined by

$$K = \{v \in H_0^1(\Omega) \mid v \geq 0 \text{ a.e.}\} = H_0^1(\Omega) \cap H_+^1(\Omega).$$

- K is no longer an affine set, but it is a *closed convex cone*:

³ Introduced by G. Stampacchia around 1963. Variational inequalities proposed by A. Signorini (1959) and studied by G. Fichera (1963) and onwards.

The Obstacle Problem³



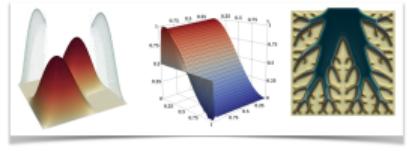
- Now minimize Dirichlet's energy over the set K defined by

$$K = \{v \in H_0^1(\Omega) \mid v \geq 0 \text{ a.e.}\} = H_0^1(\Omega) \cap H_+^1(\Omega).$$

- K is no longer an affine set, but it is a *closed convex cone*: $\alpha > 0$, $u \in K$ implies $\alpha u \in K$

³ Introduced by G. Stampacchia around 1963. Variational inequalities proposed by A. Signorini (1959) and studied by G. Fichera (1963) and onwards.

The Obstacle Problem³



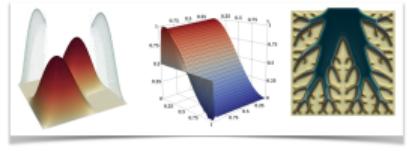
- Now minimize Dirichlet's energy over the set K defined by

$$K = \{v \in H_0^1(\Omega) \mid v \geq 0 \text{ a.e.}\} = H_0^1(\Omega) \cap H_+^1(\Omega).$$

- K is no longer an affine set, but it is a *closed convex cone*: $\alpha > 0$, $u \in K$ implies $\alpha u \in K$ and $\alpha \in (0, 1)$ and $u, v \in K$ implies $\alpha u + (1 - \alpha)v \in K$.

³ Introduced by G. Stampacchia around 1963. Variational inequalities proposed by A. Signorini (1959) and studied by G. Fichera (1963) and onwards.

The Obstacle Problem³



- Now minimize Dirichlet's energy over the set K defined by

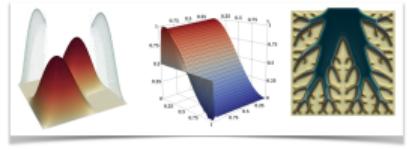
$$K = \{v \in H_0^1(\Omega) \mid v \geq 0 \text{ a.e.}\} = H_0^1(\Omega) \cap H_+^1(\Omega).$$

- K is no longer an affine set, but it is a *closed convex cone*: $\alpha > 0$, $u \in K$ implies $\alpha u \in K$ and $\alpha \in (0, 1)$ and $u, v \in K$ implies $\alpha u + (1 - \alpha)v \in K$.

We can no longer reduce the variational inequality to a system of equations,

³ Introduced by G. Stampacchia around 1963. Variational inequalities proposed by A. Signorini (1959) and studied by G. Fichera (1963) and onwards.

The Obstacle Problem³



- Now minimize Dirichlet's energy over the set K defined by

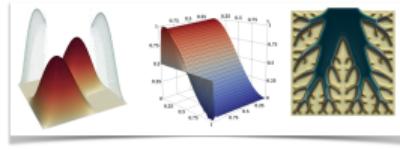
$$K = \{v \in H_0^1(\Omega) \mid v \geq 0 \text{ a.e.}\} = H_0^1(\Omega) \cap H_+^1(\Omega).$$

- K is no longer an affine set, but it is a *closed convex cone*: $\alpha > 0$, $u \in K$ implies $\alpha u \in K$ and $\alpha \in (0, 1)$ and $u, v \in K$ implies $\alpha u + (1 - \alpha)v \in K$.

We can no longer reduce the variational inequality to a system of equations, but maybe we can use the geometry to define an iterative method in which we **solve** a sequence of (nonlinear) equations?

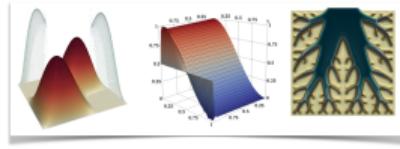
³ Introduced by G. Stampacchia around 1963. Variational inequalities proposed by A. Signorini (1959) and studied by G. Fichera (1963) and onwards.

Solving Variational Inequalities



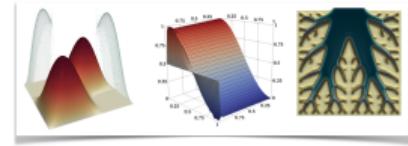
- There are many possibilities: penalty methods, interior point/barrier functions, augmented Lagrangian.

Solving Variational Inequalities



- There are many possibilities: penalty methods, interior point/barrier functions, augmented Lagrangian.
- We take an idea from convex optimization: the proximal point method and use an adaptive form of *entropy regularization*.

Solving Variational Inequalities

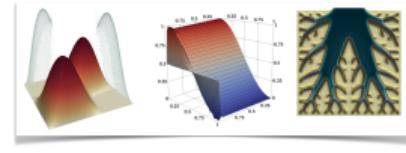


- There are many possibilities: penalty methods, interior point/barrier functions, augmented Lagrangian.
- We take an idea from convex optimization: the proximal point method and use an adaptive form of *entropy regularization*.
- Keep in mind...The obstacle problem can be viewed as a *mixed complementarity problem*: Find $(u, \lambda) \in H_0^1(\Omega) \times H^{-1}(\Omega)$ s.t.

$$\underbrace{-\Delta u - \lambda = f}_{\text{PDE}}, \quad \underbrace{u \geq 0 \quad \text{"}\lambda \geq 0\text{"}, \quad \langle \lambda, u \rangle = 0}_{\text{Complementarity}}.$$

The Lagrange multiplier λ has low regularity, this is the source of *mesh dependence* (non scalability) for many methods.

A Meta-Algorithm



Algorithm 1: Entropic proximal point algorithm for an obstacle problem.

input: Step size parameter $\alpha > 0$ and initial solution guess $w \in H_g^1(\Omega) \cap L^\infty(\Omega)$ s.t.
 $\text{ess inf } w > 0$. $f \in L^\infty(\Omega)$ and $g|_{\partial\Omega} \in C(\partial\Omega)$ s.t. $\text{ess inf}_{\partial\Omega} g > 0$.

repeat

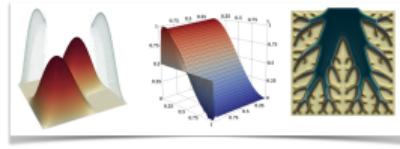
Solve the *entropic Poisson equation*,

$$\begin{cases} -\Delta u + \alpha^{-1} \ln u = f + \alpha^{-1} \ln w & \text{in } \Omega, \\ u = g & \text{on } \partial\Omega. \end{cases} \quad (3)$$

Assign $w \leftarrow u$.

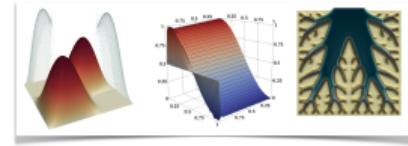
until a convergence test is satisfied

Entropic Poisson \rightarrow Saddle Point



- Formally, we can introduce **latent variables** $\psi, \tilde{\psi}$ such that $u = \exp(\psi)$ and $\tilde{\psi} = \ln w$.

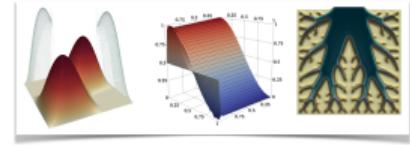
Entropic Poisson \rightarrow Saddle Point



- Formally, we can introduce **latent variables** $\psi, \tilde{\psi}$ such that $u = \exp(\psi)$ and $\tilde{\psi} = \ln w$.
- This transforms into a nonlinear saddle point problem in (u, ψ) :

$$\left\{ \begin{array}{l} \text{Find } u \in H_g^1(\Omega) \text{ and } \psi \in W \text{ such that} \\ \int_{\Omega} \alpha \nabla u \cdot \nabla v dx + \int_{\Omega} \psi v dx = \int_{\Omega} (\alpha f + w) v dx \text{ for all } v \in H_0^1(\Omega), \\ \int_{\Omega} u \varphi dx - \int_{\Omega} \exp(\psi) \varphi dx = 0 \quad \text{for all } \varphi \in U. \end{array} \right.$$

Entropic Poisson \rightarrow Saddle Point

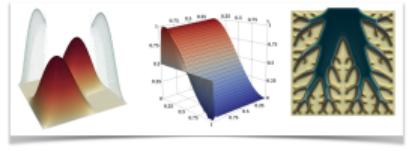


- Formally, we can introduce **latent variables** $\psi, \tilde{\psi}$ such that $u = \exp(\psi)$ and $\tilde{\psi} = \ln w$.
- This transforms into a nonlinear saddle point problem in (u, ψ) :

$$\left\{ \begin{array}{l} \text{Find } u \in H_g^1(\Omega) \text{ and } \psi \in W \text{ such that} \\ \int_{\Omega} \alpha \nabla u \cdot \nabla v dx + \int_{\Omega} \psi v dx = \int_{\Omega} (\alpha f + w) v dx \text{ for all } v \in H_0^1(\Omega), \\ \int_{\Omega} u \varphi dx - \int_{\Omega} \exp(\psi) \varphi dx = 0 \quad \text{for all } \varphi \in U. \end{array} \right.$$

Regardless of the choice or order of approximating spaces for $H_g^1(\Omega)$, W , and U , the discrete solution ψ_h yields a latent solution $\tilde{u}_h := \exp(\psi_h)$ that is globally feasible on the discrete level.

Proximal Galerkin



Algorithm 2:

Input : Step size parameter $\alpha > 0$, linear subspaces $V_h \subset H_0^1(\Omega)$ and $W_h \subset L^\infty(\Omega)$, and initial solution guess $\psi_h \in W_h$.

Output: Approximate solutions u_h and $\tilde{u}_h = \exp \psi_h$, and approximate Lagrange multiplier, $\lambda_h = (\omega_h - \psi_h)/\alpha$.

repeat

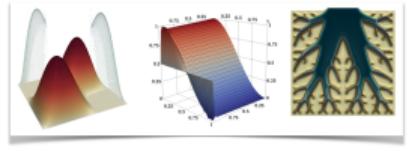
 Assign $\omega_h \leftarrow \psi_h$.

 Solve the following (nonlinear) discrete saddle-point problem:

$$\begin{cases} \text{Find } u_h \in g_h + V_h \text{ and } \psi_h \in W_h \text{ such that} \\ \int_{\Omega} \alpha \nabla u_h \cdot \nabla v dx + \int_{\Omega} \psi_h v dx = \int_{\Omega} (f + \omega_h) v dx \text{ for all } v \in V_h, \\ \int_{\Omega} u_h \varphi dx - \int_{\Omega} \exp(\psi_h) \varphi dx = 0 \quad \text{for all } \varphi \in U_h. \end{cases}$$

until a convergence test is satisfied

Proximal Galerkin



Algorithm 3:

Input : Step size parameter $\alpha > 0$, linear subspaces $V_h \subset H_0^1(\Omega)$ and $W_h \subset L^\infty(\Omega)$, and initial solution guess $\psi_h \in W_h$.

Output: Approximate solutions u_h and $\tilde{u}_h = \exp \psi_h$, and approximate Lagrange multiplier, $\lambda_h = (\omega_h - \psi_h)/\alpha$.

repeat

 Assign $\omega_h \leftarrow \psi_h$.

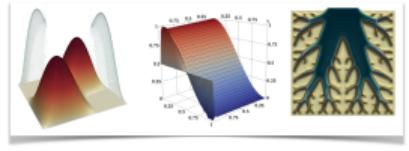
 Solve the following (nonlinear) discrete saddle-point problem:

$$\begin{cases} \text{Find } u_h \in g_h + V_h \text{ and } \psi_h \in W_h \text{ such that} \\ \int_{\Omega} \alpha \nabla u_h \cdot \nabla v dx + \int_{\Omega} \psi_h v dx = \int_{\Omega} (f + \omega_h) v dx \text{ for all } v \in V_h, \\ \int_{\Omega} u_h \varphi dx - \int_{\Omega} \exp(\psi_h) \varphi dx = 0 \quad \text{for all } \varphi \in U_h. \end{cases}$$

until a convergence test is satisfied

- The *full algorithm* involves updating α , ω and repeating Algorithm 2.
- $\alpha > 0$ can remain fixed or updated successively provided $\sum \alpha = \infty$.
- Convergence of this *outer loop* is provided (on an ∞ -dimensional level) by the classical proximal point method with a rate $O((\sum \alpha)^{-1/2})$.

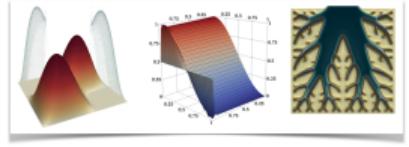
Finite Element Spaces I



- \mathcal{T}_h shape-regular partition of $\Omega \subset \mathbb{R}^2$.
- $T \in \mathcal{T}_h$ open connected triangular mesh cells with Lipschitz boundaries ∂T .
- $\Omega := \bigcup_{T \in \mathcal{T}_h} \overline{T}$.
- $h = \max_{T \in \mathcal{T}_h} \text{diam}(T)$ is the mesh size.
- $\mathbb{P}_p(T)$ space of polynomials of total order up to and including p on a triangle T .
- $\mathbb{X}(T)$ of polynomials over an element $T \in \mathcal{T}_h$.
- We define “broken” polynomials

$$\mathbb{X}(\mathcal{T}_h) = \{\varphi \in L^\infty(\Omega) \mid \varphi|_T \in \mathbb{X}(T) \text{ for every } T \in \mathcal{T}_h\}.$$

Finite Element Spaces II



- We require spaces of degree- q polynomials on whose traces on the cell boundary ∂T have lower polynomial degree $p < q$.
- Define the sets of bubble functions in $\mathbb{P}_q(T)$ and $\mathbb{Q}_q(T)$ to be

$$\mathring{\mathbb{P}}^q(T) = \{\varphi \in \mathbb{P}_q(T) \mid \varphi|_{\partial T} = 0\}$$

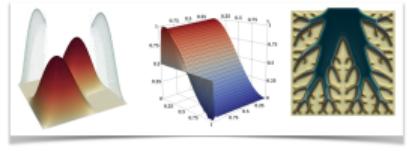
- Then define

$$\hat{\mathbb{P}}_p(T) = \mathbb{P}_p(T) \setminus \mathring{\mathbb{P}}^p(T)$$

- Finally let

$$\mathbb{P}_p^q(T) = \hat{\mathbb{P}}_p(T) \oplus \mathring{\mathbb{P}}^q(T). \quad (4)$$

Finite Element Spaces III

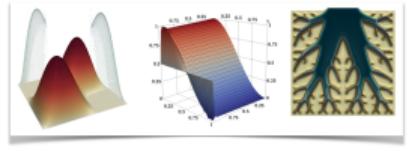


For any integer $p \geq 1$, we define the following pairs of spaces: We refer to the following as the $(\mathbb{P}_p\text{-bubble}, \mathbb{P}_{p-1}\text{-broken})$ pairing:

$$V_h = \mathbb{P}_p^{p+2}(\mathcal{T}_h) \cap H_0^1(\Omega), \quad W_h = \mathbb{P}_{p-1}(\mathcal{T}_h).$$

- Example, $p = 1$: V_h is composed of the direct sum of continuous piecewise linear functions and 3rd order bubble functions and W_h is piecewise constants.

Finite Element Spaces III



For any integer $p \geq 1$, we define the following pairs of spaces: We refer to the following as the $(\mathbb{P}_p\text{-bubble}, \mathbb{P}_{p-1}\text{-broken})$ pairing:

$$V_h = \mathbb{P}_p^{p+2}(\mathcal{T}_h) \cap H_0^1(\Omega), \quad W_h = \mathbb{P}_{p-1}(\mathcal{T}_h).$$

- Example, $p = 1$: V_h is composed of the direct sum of continuous piecewise linear functions and 3rd order bubble functions and W_h is piecewise constants.
- For shape regular \mathcal{T}_h , these pairs of spaces are **stable** for the linearized, singularly perturbed saddle point problems.
- The paper contains similar spaces for quadrilateral mesh cells and alternative pairs without bubble functions that are also stable.

Discrete Domains

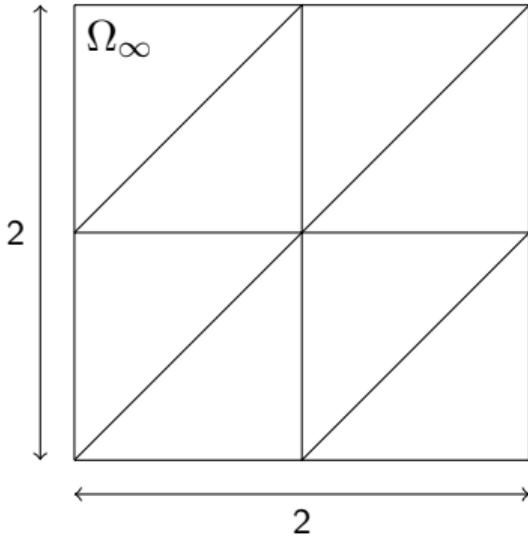
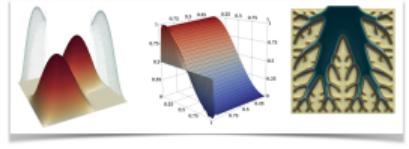
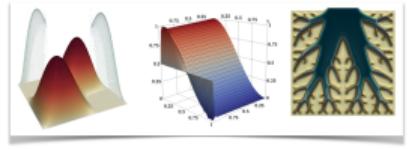


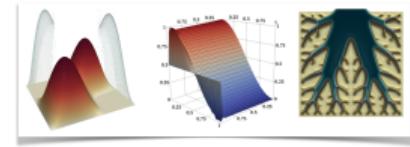
Figure: Initial FE meshes for domain Ω_∞ mesh sizes $h = h_\infty$, respectively. (I): Initial triangular mesh for $(\mathbb{P}_p\text{-bubble}, \mathbb{P}_{p-1}\text{-broken})$ pair on Ω_∞ . We consider various polynomial orders $p \geq 1$ and mesh sizes.

Experiment 1



How does PG behave over various meshes and polynomials orders?

Experiment 1



- Set $g = u$, where $u(x, y)$ is the smooth manufactured solution

$$u(x, y) = \begin{cases} 0 & \text{if } x < 0, \\ x^4 & \text{otherwise,} \end{cases} \quad \text{implied by} \quad f(x, y) = \begin{cases} 0 & \text{if } x < 0, \\ -12x^2 & \text{otherwise.} \end{cases} \quad (5)$$

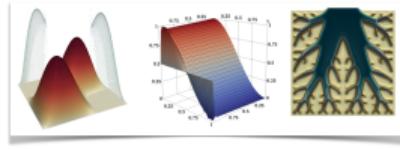
- $\lambda \equiv 0$: this is a **biactive** solution, meaning problem is nonsmooth at the solution.
- Proximal point is a (slow) fixed point method if α is left fixed. We choose:

$$\alpha_1 = 1, \quad \alpha_k = \min\{\max\{\alpha_1, r^{q^{k-1}} - \alpha_{k-1}\}, 10^{10}\}, \quad k = 2, 3, \dots, \quad (6)$$

where $r = q = 1.5$.

- Once $\alpha_k = 10^{10}$, we can check successive iterates as a stopping criterion.

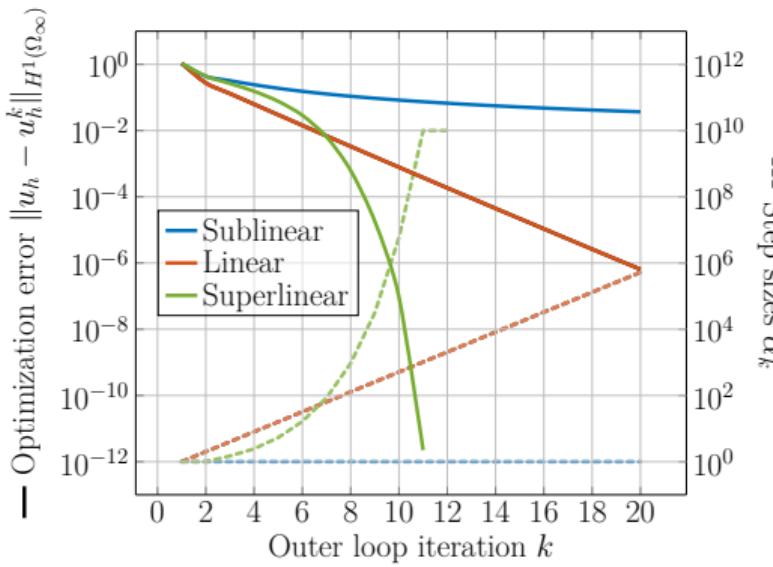
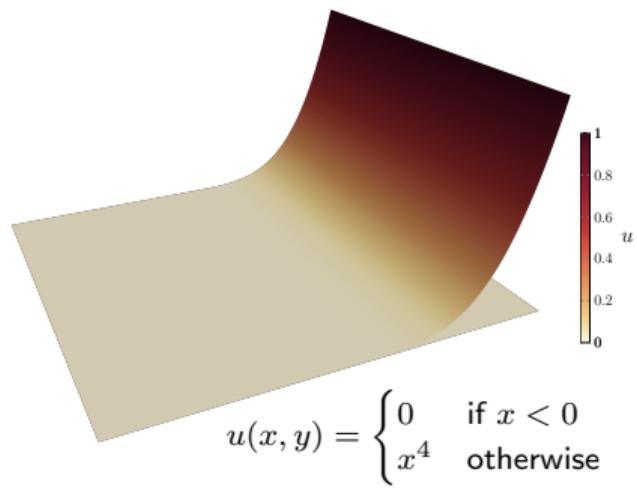
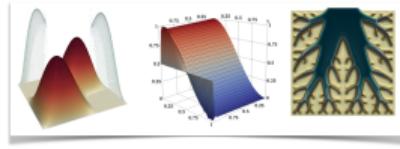
Experiment 1: Results



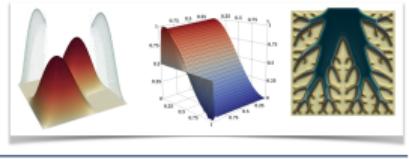
Progress of the iterates $\|u_h^k - u_h^{k-1}\|_{H^1(\Omega_\infty)}$ for various h and p

k	α_k	Polynomial order $p = 1$			Polynomial order $p = 2$	
		$h_\infty/16$	$h_\infty/32$	$h_\infty/64$	$h_\infty/16$	$h_\infty/32$
1	1.0	$2.10 \cdot 10^0$	$2.10 \cdot 10^0$	$2.10 \cdot 10^0$	$2.10 \cdot 10^0$	$2.10 \cdot 10^0$
2	1.0	$6.45 \cdot 10^{-1}$	$6.45 \cdot 10^{-1}$	$6.45 \cdot 10^{-1}$	$6.45 \cdot 10^{-1}$	$6.45 \cdot 10^{-1}$
3	1.49	$1.73 \cdot 10^{-1}$	$1.73 \cdot 10^{-1}$	$1.73 \cdot 10^{-1}$	$1.73 \cdot 10^{-1}$	$1.73 \cdot 10^{-1}$
4	2.43	$1.10 \cdot 10^{-1}$	$1.10 \cdot 10^{-1}$	$1.10 \cdot 10^{-1}$	$1.10 \cdot 10^{-1}$	$1.10 \cdot 10^{-1}$
5	5.35	$7.76 \cdot 10^{-2}$	$7.77 \cdot 10^{-2}$	$7.77 \cdot 10^{-2}$	$7.77 \cdot 10^{-2}$	$7.77 \cdot 10^{-2}$
6	$1.64 \cdot 10^1$	$4.76 \cdot 10^{-2}$	$4.77 \cdot 10^{-2}$	$4.77 \cdot 10^{-2}$	$4.77 \cdot 10^{-2}$	$4.77 \cdot 10^{-2}$
7	$8.50 \cdot 10^1$	$2.24 \cdot 10^{-2}$	$2.25 \cdot 10^{-2}$	$2.25 \cdot 10^{-2}$	$2.25 \cdot 10^{-2}$	$2.25 \cdot 10^{-2}$
8	$9.35 \cdot 10^2$	$5.82 \cdot 10^{-3}$	$5.84 \cdot 10^{-3}$	$5.85 \cdot 10^{-3}$	$5.85 \cdot 10^{-3}$	$5.85 \cdot 10^{-3}$
9	$3.17 \cdot 10^4$	$6.04 \cdot 10^{-4}$	$6.07 \cdot 10^{-4}$	$6.07 \cdot 10^{-4}$	$6.07 \cdot 10^{-4}$	$6.07 \cdot 10^{-4}$
10	$5.85 \cdot 10^6$	$1.80 \cdot 10^{-5}$	$1.81 \cdot 10^{-5}$	$1.81 \cdot 10^{-5}$	$1.81 \cdot 10^{-5}$	$1.81 \cdot 10^{-5}$
11	$1 \cdot 10^{10}$	$9.41 \cdot 10^{-8}$	$9.47 \cdot 10^{-8}$	$9.49 \cdot 10^{-8}$	$9.50 \cdot 10^{-8}$	$9.50 \cdot 10^{-8}$
12	$1 \cdot 10^{10}$	$2.10 \cdot 10^{-12}$	$2.00 \cdot 10^{-12}$	$1.96 \cdot 10^{-12}$	$1.92 \cdot 10^{-12}$	$1.95 \cdot 10^{-10}$
Tot. linear solves		21	20	19	19	19

Experiment 1: Results

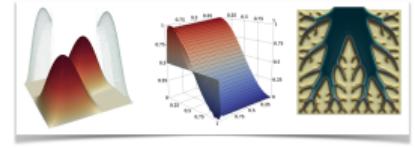


Experiment 2



How does PG on optimality conditions and strict complementarity?

Experiment 2



- In this experiment, we set $g = 0$ and define

$$f(x, y) = 2\pi^2 \sin(\pi x) \sin(\pi y). \quad (7)$$

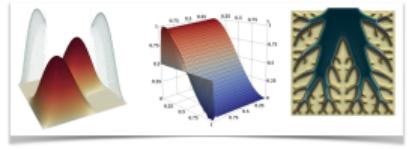
- We aim to check convergence of the discrete solution via the KKT conditions:

$$\left| \int_{\Omega} \lambda u dx \right| = 0, \text{ complementary slackness,}$$

$$\int_{\Omega} \max\{-u, 0\} dx = 0, \text{ primal feasibility,}$$

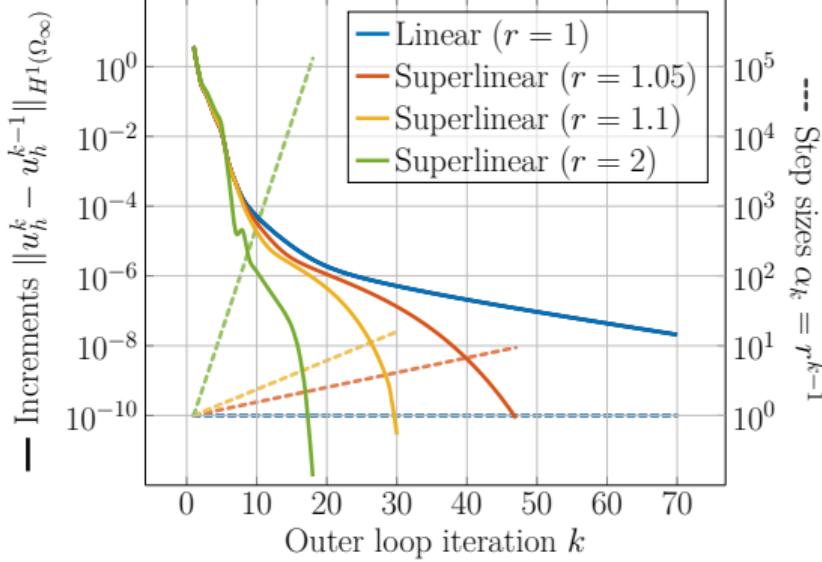
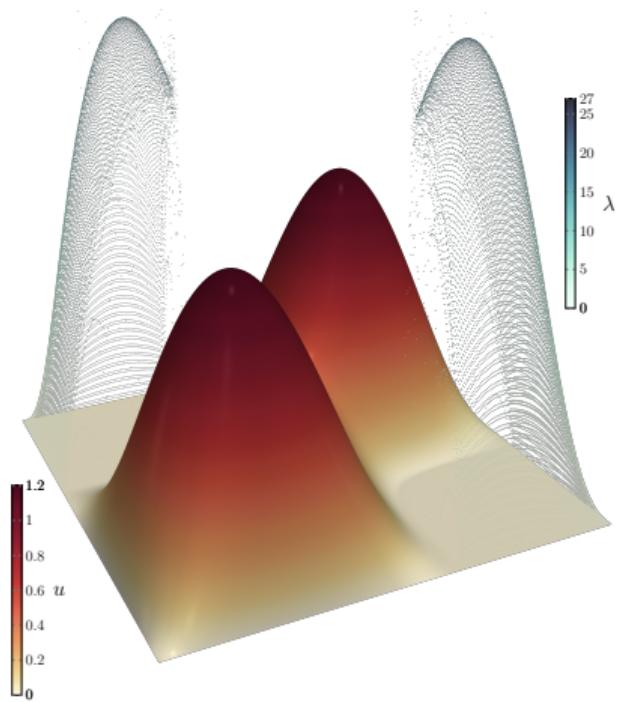
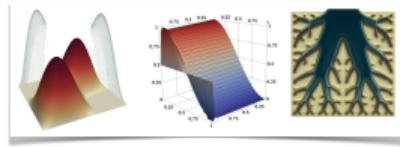
$$\int_{\Omega} \max\{-\lambda, 0\} dx = 0, \text{ dual feasibility.}$$

Experiment 2: Results

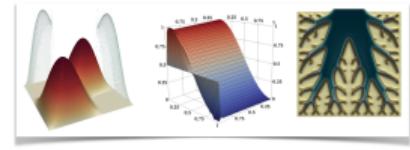


h	Complementarity $\left \int_{\Omega_\infty} \lambda_h u_h dx \right $	Primal feasibility $\int_{\Omega_\infty} \max\{-u_h, 0\} dx$	Dual feasibility $\int_{\Omega_\infty} \max\{-\lambda_h, 0\} dx$
h_∞		$6.97 \cdot 10^{-3}$	
$h_\infty/2$		$9.09 \cdot 10^{-3}$	
$h_\infty/4$	(all less than 10^{-14})	$1.16 \cdot 10^{-3}$	(all less than 10^{-12})
$h_\infty/8$		$1.69 \cdot 10^{-4}$	
$h_\infty/16$		$4.08 \cdot 10^{-5}$	
$h_\infty/32$		$4.53 \cdot 10^{-6}$	

Experiment 2: Results

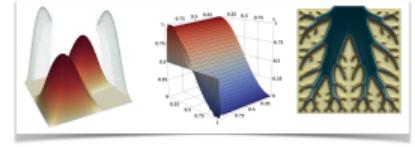


Experiment 3



How does PG do with nontrivial obstacles and higher order spaces?

Experiment 3



- Set $f = 0$, $g = 0$ and define the obstacle to be the upper surface of a sphere of radius $1/2$, namely

$$\phi(x, y) = \sqrt{1/4 - x^2 - y^2}, \quad (8)$$

if $\sqrt{x^2 + y^2} \leq 1/2$, and assume that ϕ is sufficiently negative when $\sqrt{x^2 + y^2} > 1/2$ so that no contact happens on that subdomain.

- The exact solution on the circular domain $\Omega = \Omega_2$ is found to be

$$u(x, y) = \begin{cases} A \ln \sqrt{x^2 + y^2} & \text{if } \sqrt{x^2 + y^2} > a, \\ \phi(x, y) & \text{otherwise,} \end{cases} \quad (9)$$

where $a = \exp(W_{-1}(-1/(2e^2))/2 + 1) \approx 0.34898$, $A = \sqrt{1/4 - a^2}/\ln a \approx -0.34012$, and $W_j(\cdot)$ is the j -th branch of the Lambert W-function.

Experiment 3: Results

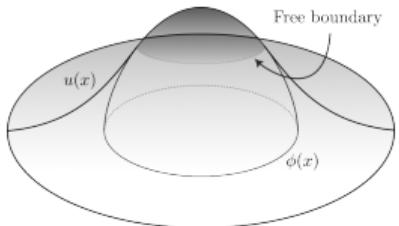
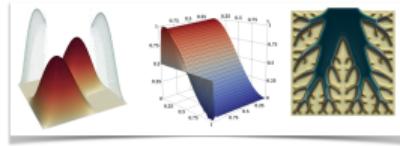
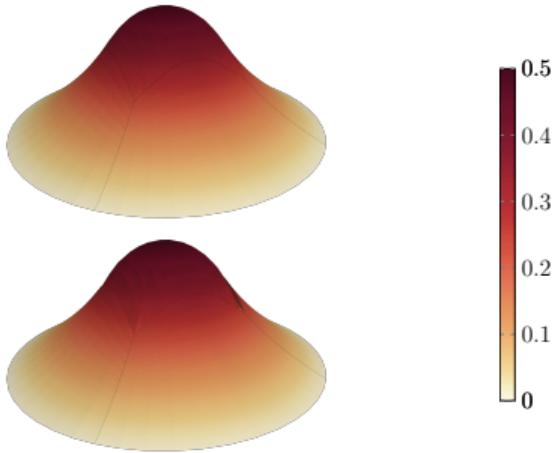


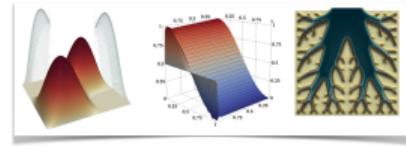
Diagram of problem



Very high order ($p = 12$) proximal Galerkin
solutions u_h (top) and \tilde{u}_h (bottom)

Figure: Spherical obstacle. Benchmark obstacle problem. Left: Diagram of the problem set-up. Right: Five-element proximal Galerkin solutions u_h and \tilde{u}_h .

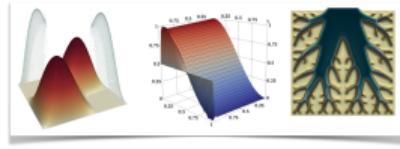
Experiment 3: Results



		Primal errors $\ u - u_h^k\ _{H^1(\Omega_2)}$ for $p = 1$				
k	Linear solves	$h_2/8$	$h_2/16$	$h_2/32$	$h_2/64$	$h_2/128$
1	3	$2.72 \cdot 10^{-1}$	$2.70 \cdot 10^{-1}$	$2.70 \cdot 10^{-1}$	$2.70 \cdot 10^{-1}$	$2.70 \cdot 10^{-1}$
2	1	$1.37 \cdot 10^{-1}$	$1.38 \cdot 10^{-1}$	$1.38 \cdot 10^{-1}$	$1.38 \cdot 10^{-1}$	$1.38 \cdot 10^{-1}$
3	1	$3.62 \cdot 10^{-2}$	$3.33 \cdot 10^{-2}$	$3.31 \cdot 10^{-2}$	$3.31 \cdot 10^{-2}$	$3.31 \cdot 10^{-2}$
:	:	:	:	:	:	:
Total iterations		11	11	11	11	11
Total linear solves		13	13	13	13	13
Final error		$1.98 \cdot 10^{-2}$	$8.73 \cdot 10^{-3}$	$3.49 \cdot 10^{-3}$	$1.18 \cdot 10^{-3}$	$3.85 \cdot 10^{-4}$

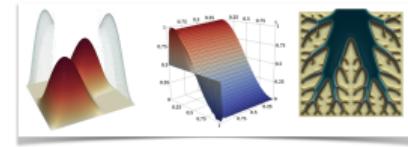
Table: Subproblem error, $\|u - u_h^k\|_{H^1(\Omega_2)}$, for various mesh sizes using $(\mathbb{Q}_2, \mathbb{Q}_0\text{-broken})$ discretization. We used $\alpha_k = 1$ for all $k = 1, \dots$ and stopped the algorithm when $\|u_h^k - u_h^{k-1}\|_{L^2(\Omega_2)} < 10^{-6}$.

Nonsymmetric VI



What about variational inequalities that do not arise as energy minimization principles?

Ext. 1: Discrete Maximum Principle

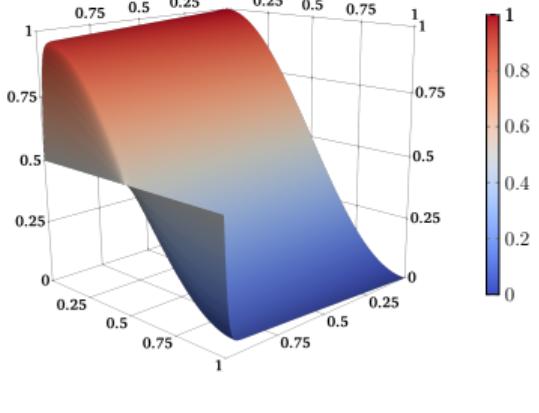
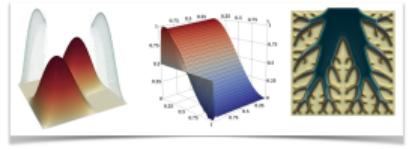


- Let $\epsilon > 0$, $\beta \in \mathbb{R}^d$ be fixed, $f \in L^\infty(\Omega)$, and $g \in H^1(\Omega) \cap C(\bar{\Omega})$.
- The solution u to the advection diffusion equation should satisfy a maximum principle on the continuous and **discrete** level:

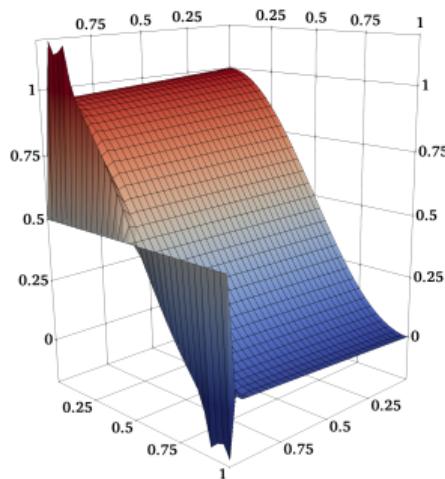
$$-\epsilon \Delta u + \beta \cdot \nabla u = f \quad \text{in } \Omega, \quad u = g \quad \text{on } \partial\Omega, \quad (10)$$

- By rewriting (10) as a variational inequality, we can develop a Proximal Galerkin scheme.
- There are (many) details in the paper including a “binary entropic Poisson equation”, an implementable algorithm, further pairs of tailored FE spaces for the new saddle point problems.

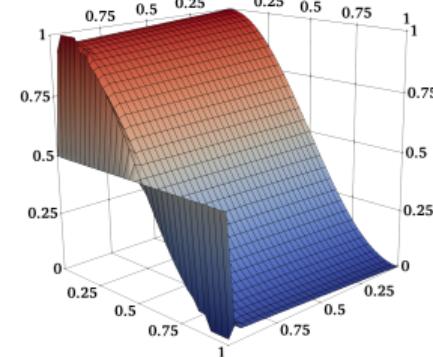
Ext. 1: Discrete Maximum Principle



Exact solution u



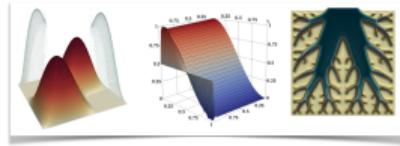
FEM solution



Proximal Galerkin solution \tilde{u}_h

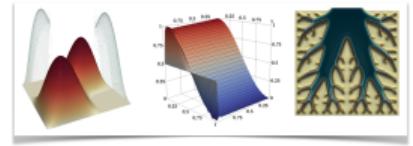
Figure: Eriksson–Johnson problem with $\epsilon = 10^{-2}$. (l): Exact solution. (c): A first-order Bubnov–Galerkin numerical solution that clearly violates the strong maximum principle $0 \leq u(x) \leq 1$. (r): $(\mathbb{Q}_1, \mathbb{Q}_1)$ -proximal Galerkin solution $\tilde{u}_h = \text{expit}(\psi_h)$ satisfies the strong maximum principle, by construction.

Nonconvex Problems



What about nonconvex variational problems?

Ext. 2.: Topology Optimization



Find a material distribution with maximum flexibility and a fixed volume:

$$\min_{\rho \in L^\infty(\Omega)} \left\{ \widehat{F}(\mathbf{u}, \rho) = \int_{\Omega} \mathbf{u} \cdot \mathbf{f} dx \right\}, \quad (11)$$

subject to the constraints

$$\begin{cases} -\operatorname{Div}(r(\tilde{\rho}) \boldsymbol{\sigma}) = \mathbf{f} \text{ in } \Omega \text{ with } \mathbf{u} = 0 \text{ on } \Gamma_0, \quad \boldsymbol{\sigma} \mathbf{n} = 0 \text{ on } \partial\Omega \setminus \Gamma_0, \\ -\epsilon^2 \Delta \tilde{\rho} + \tilde{\rho} = \rho \text{ in } \Omega \text{ with } \nabla \tilde{\rho} \cdot \mathbf{n} = 0 \text{ on } \partial\Omega, \\ \int_{\Omega} \rho(\mathbf{x}) d\mathbf{x} = \theta |\Omega|, \quad 0 \leq \rho \leq 1, \quad r(\tilde{\rho}) = \underline{\rho} + \tilde{\rho}^3(1 - \underline{\rho}), \end{cases} \quad (12)$$

where $\epsilon > 0$ is a *length scale* and $0 < \theta < 1$ is the desired volume fraction, which constrains the amount of the domain Ω occupied by the design.

Ext. 2: Topology Optimization

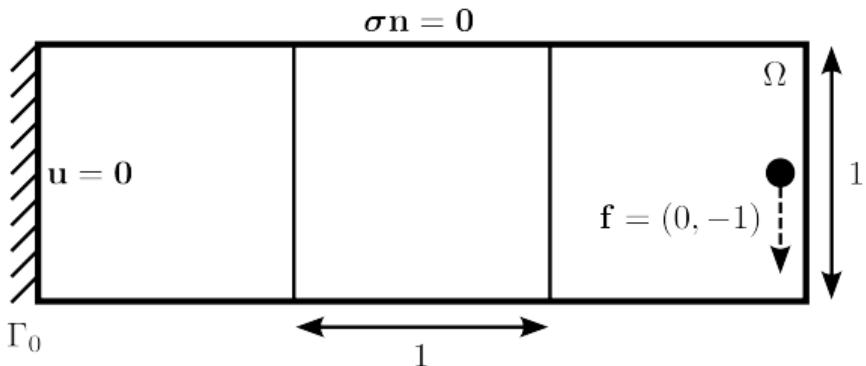
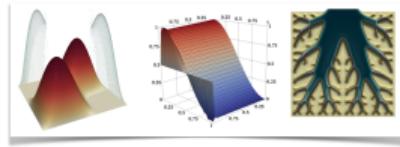
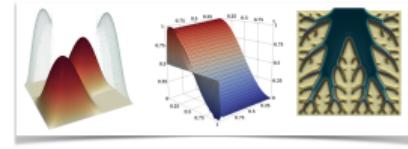


Figure: The design domain Ω for the cantilever beam problem with corresponding boundary conditions and three-element initial mesh with length $h_0 = 1$.

Mirror Descent



- Like proximal point, but replace the objective with its linearization.

Algorithm 4: Entropic mirror descent for topology optimization.

Input : Initial latent variable $\rho^0 \in L^\infty(\Omega)$, sequence of step sizes $\alpha_k > 0$, increment tolerance $\text{itol.} > 0$, and normalized tolerance $\text{ntol.} > 0$.
Output: Optimized material density $\bar{\rho} = \text{expit}(\psi^k)$.

Initialize $k = 0$.

while $\|\text{expit}(\psi^k) - \text{expit}(\psi^{k-1})\|_{L^1(\Omega)} > \min\{\alpha_k \text{ntol.}, \text{itol.}\}$ **do**

// Latent space gradient descent

Assign $\psi^{k+1/2} \leftarrow \psi^k - \alpha_{k+1} \nabla F(\text{expit}(\psi^k))$.

// Compute Lagrange multiplier

Solve for $c \in \mathbb{R}$ such that $\int_{\Omega} \text{expit}(\psi^{k+1/2} + c) dx = \theta |\Omega|$.

// Latent space feasibility correction

Assign $\psi^{k+1} \leftarrow \psi^{k+1/2} + c$.

Assign $k \leftarrow k + 1$.

Ext. 2: Results

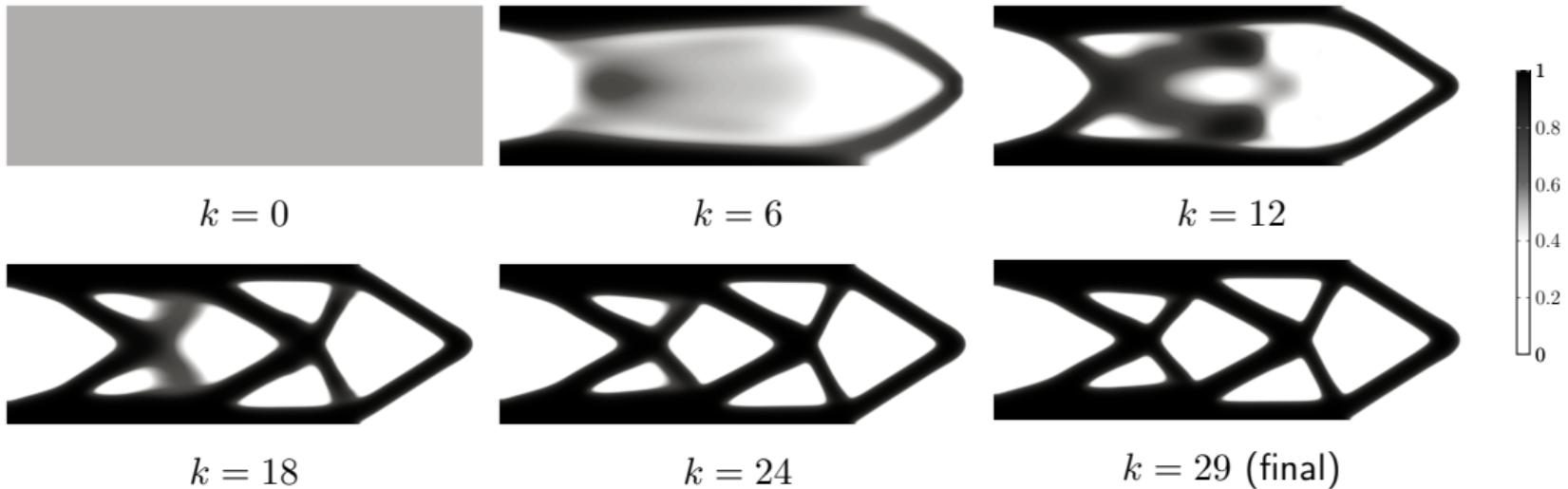
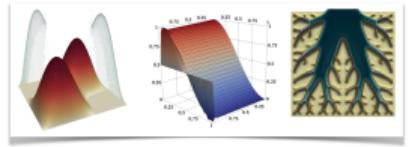
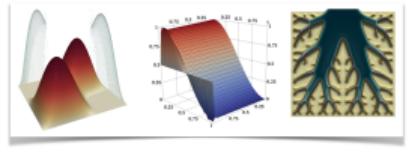


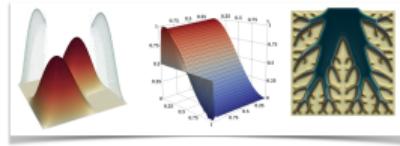
Figure: Subsequence of material densities $\tilde{\rho}_h^k$ from Algorithm 4 for selected iterations k . Results obtained with problem parameters $\epsilon = 2 \cdot 10^{-2}$ and $\theta = 0.5$; algorithm parameters $\text{itol.} = 10^{-2}$, $\text{ntol.} = 10^{-5}$, and $\alpha_k = 25k$; and discretization parameters $h = h_0/128$ and $p = 1$.

Latent Variable Proximal Point



What is really going on here?

A



The LVPP Method: We seek a minimizer \bar{u} of a smooth coercive functional J over a closed convex subset K of a Sobolev space V . We assume $K = \{v \in V \mid \Lambda v \in C \text{ a.e. } \Omega\}$, Ω is an open bounded set, Λ is a bounded linear operator into $L^2(\Omega)^n$ or $L^2(\Omega)^{n \times n}$, and C is a nonempty closed convex subset of \mathbb{R}^n or $\mathbb{R}^{n \times n}$. This enforces properties like non-negativity or boundedness of Λv almost everywhere on Ω .

Just as most PDEs do not admit analytical solutions in general, we require a numerical algorithm to approximate \bar{u} . The ideal method would be provably convergent, accurate, and exhibit low iteration complexity. Here, *accuracy* refers to an achievable tolerance and *complexity* to the number of steps required until a stopping criterion is fulfilled at that tolerance. These infinite-dimensional problems have the added challenge that V must be replaced by a finite-dimensional space V_h in practice. This adds another desirable property: *mesh-independence*, i.e. the iteration complexity for a fixed accuracy is independent of the discretization parameter h . Mesh-independence is deeply linked to the properties of V, J, K , and \bar{u} itself at the *continuous level* [?] and entails an intricate analysis for constrained problems [?]. For this reason, many researchers design and