SpezialForschungsBereich F 32

Karl–Franzens Universität Graz
Technische Universität Graz
Medizinische Universität Graz

A discretization method for the numerical solution of Dieudonné-Rashevsky type problems with application to edge detection within noisy image data

Lucas Franek Marzena Franek
Helmut Maurer Marcus Wagner

SFB-Report No. 2010–030 August 2010

A–8010 GRAZ, HEINRICHSTRASSE 36, AUSTRIA

Supported by the
Austrian Science Fund (FWF)

SFB sponsors:

- **Austrian Science Fund** (FWF)
- **University of Graz**
- **Graz University of Technology**
- **Medical University of Graz**
- **Government of Styria**
- **City of Graz**

A discretization method for the numerical solution of Dieudonné-Rashevsky type problems with application to edge detection within noisy image data

Lucas Franek, Marzena Franek, Helmut Maurer and Marcus Wagner

1. Introduction.

The present paper is concerned with the numerical solution of multidimensional control problems of the following type:

$$(\mathrm{P}): \quad F(x,u) = \int_\Omega f(s,x(s),u(s))\,ds \longrightarrow \mathrm{Min}\,!\,; \quad (x,u) \in W_0^{1,p}(\Omega,\mathbb{R}^n) \times L^p(\Omega,\mathbb{R}^{nm})\,; \tag{1.1}$$

$$Jx(s) = \begin{pmatrix} \dfrac{\partial x_1}{\partial s_1}(s) & \ldots & \dfrac{\partial x_1}{\partial s_m}(s) \\ \vdots & & \vdots \\ \dfrac{\partial x_n}{\partial s_1}(s) & \ldots & \dfrac{\partial x_n}{\partial s_m}(s) \end{pmatrix} = B(s)\,u(s) \;\; (\forall)\, s \in \Omega\,; \tag{1.2}$$

$$u \in \mathrm{U} = \left\{ u \in L^p(\Omega,\mathbb{R}^{nm}) \;\middle|\; u(s) \in \mathrm{G} \;\; (\forall)\, s \in \Omega \right\}. \tag{1.3}$$

Problems of this kind, referred as Dieudonné-Rashevsky type problems, are connected with the study of BVP's for nonlinear first-order PDE's, [01] they arise from optimization problems for convex bodies under geometrical restrictions, [02] in elasticity theory (torsion problems) [03] and population dynamics (models with comprehension of an age structure). [04] Recently, problems from image processing have been studied under consideration of gradient constraints as well (see below).

Although Dieudonné-Rashevsky type problems have been as yet less intensely investigated than optimal control problems with second-order PDE's, [05] necessary optimality conditions as well as duality theorems have been derived in the linear-convex case (with state equations $Jx(s) - u(s) = 0$ resp. $Jx(s) - A(s)\,x(s) - B(s)\,u(s) = 0$). [06] A numerical approach, however, has been first proposed in the context of applications from image processing as a direct method: [07] the problem (P) will be first discretized, and the arising large-scale finite-dimensional problems will be numerically solved by interior-point methods. In the present paper, we first prove a convergence theorem for this discretization method (*Section 2*). By use of the global minimizers of the discretized problems, we construct a minimizing sequence $\{\,(\widetilde{x}_N, \widetilde{u}_N)\,\}$ of admissible solutions of (P)

---

01) [Dacorogna/Marcellini 97], [Dacorogna/Marcellini 98] and [Dacorogna/Marcellini 99].

02) [Andrejewa/Klötzler 84a] and [Andrejewa/Klötzler 84b], p. 149 f.

03) [Funk 62], pp. 531 ff., [Lur'e 75], pp. 240 ff., [Ting 69a], p. 531 f., [Ting 69b] and [Wagner 96], pp. 76 ff.

04) [Brokate 85], [Feichtinger/Tragler/Veliov 03].

05) See e. g. [Tröltzsch 05].

06) The corresponding theory has been developed substantially by Klötzler, Pickenhain and Wagner, cf. the bibliography in [Wagner 09], pp. 545 and 548.

07) [Brune/Maurer/Wagner 09], [Franek 07a], [Franek 07b]. The only precursor is [Dewess/Helbig 95] where a transportation flow problem as the *dual* problem to a Dieudonné-Rashevsky type problem has been solved by methods of combinatorial optimization.

with $|F(\widetilde{x}_N, \widetilde{u}_N) - m| \leqslant C\,\sigma_N$ where m and σ_N denote the minimal value of (P) and the mesh size of the triangulation in the discretized problem, respectively. Within this minimizing sequence, there exists a subsequence, which converges uniformly w. r. to x and L^p-weakly ($1 < p < \infty$) w. r. to u.

In *Section 3*, we present an application of our optimal control method to a basic problem from image processing, namely to the problem of edge detection within noisy image data. There are numerous situations where the introduction of gradient constraints into problems of image processing is highly reasonable. In the according reformulation of the image denoising problem and the optical flow problem, the optimal control approach allows a simultaneous edge determination as the subregions of Ω where the control restriction becomes (nearly) active. [08] In the Shape from Shading problem, the solutions of the underlying first-order PDE (Horn's equation) obey from the outset a convex gradient constraint, which has to be considered explicitly after the variational reformulation of the problem, thus converting the resulting multidimensional variational problem into an optimal control problem of type (P). [09] In the elastic image registration problem, where a correspondence between a template and a reference image is searched as an elastic deformation x, the incorporation of gradient constraints ensures the validity of the underlying elasticity model since the derivatives of x correspond with the shear stresses which should remain bounded in the linear-elastic as well as in the hyperelastic case. [10]

Our choice for the study of the problem of edge detection within noisy image has different motivations. Firstly, the *application of our newly proposed method to a well-examined problem* is desirable for methodical reasons. Indeed, the problem of edge detection has been accessed by a number of different approaches. If considered as a generalization of the classical image denoising problem, it has been solved by variational methods (see Section 3.b) below), [11] but in the context of image segmentation, it has been treated by level-set methods [12] and graph-theoretical methods [13] as well. Our second purpose is to perform a *visual as well as a quantitative comparison* of the results from the optimal control approach with the output of other methods. In the present paper, we focus to the comparison with edge sketches generated by EDISON, a well-introduced non-variational method, [14] as well as by a state-of-art variational method, the edge detection via Ambrosio-Tortorelli type functional. The reference to variational methods is particularly advisable since the optimal control formulation appears as their natural generalization. We may summarize that the quality of the results given by our optimal control method competes well with those given by the two selected reference methods. When applied to noisy data, the optimal control method exceeds the Ambrosio-Tortorelli method in quality in all cases, and in some cases even the sketches generated by EDISON (see Sections 3.e) and 3.f) below). In comparison with the Ambrosio-Tortorelli method, the optimal control approach provides some advantages even with respect to implementation and runtime behaviour.

A third purpose is to work out the optimal control approach as a "building block" for the treatment of situations where the *edge detection within noisy image data must be embedded as a subtask* into a more comprehensive problem. In this respect, we refer particularly to the problem of multimodal image matching. Here the registration of the images has to be performed in absence of a greyscale value correspondence,

---

[08] [Brune/Maurer/Wagner 09], [Wagner 06], pp. 108 ff., [Wagner 09], pp. 558 ff.

[09] See [Wagner 09], pp. 564 ff., and the references cited therein.

[10] [Wagner 08], pp. 26 ff., [Wagner 10].

[11] Cf. the overviews in [Aubert/Kornprobst 06], pp. 65 ff., and [Buades/Coll/Morel 05].

[12] [Aubert/Kornprobst 06], pp. 173 ff., [Chan/Esedoḡlu/Nikolova 06], [Chan/Vese 01].

[13] [Felzenszwalb/Huttenlocher 04] and [Shi/Malik 00].

[14] Cf. [Meer/Georgescu 01]. The program is accessible via http://www.caip.rutgers.edu/riul/research/code/EDISON/edison_binary.zip (last access: 21.12.2009).

consequently, the matching must be based exclusively on the geometrical information contained in the images, which will be represented as gradient or edge information, respectively. [15)]

Notations.

Assume that $\Omega \subset \mathbb{R}^m$ is the closure of a bounded strongly Lipschitz domain. Then $C^k(\Omega, \mathbb{R}^r)$ denotes the space of r-dimensional vector functions $f\colon \Omega \to \mathbb{R}^r$, whose components are continuous ($k = 0$) resp. k-times continuously differentiable ($k = 1, \ldots, \infty$). $L^p(\Omega, \mathbb{R}^r)$ denotes the space of r-dimensional vector functions $f\colon \Omega \to \mathbb{R}^r$ with components, which are integrable in the pth power ($1 \leqslant p < \infty$) resp. are measurable and essentially bounded ($p = \infty$). $W_0^{1,p}(\Omega, \mathbb{R}^r)$ denotes the Sobolev space of compactly supported $L^p(\Omega, \mathbb{R}^r)$ (vector) functions $f\colon \Omega \to \mathbb{R}^r$, whose components possess first-order weak partial derivatives in $L^p(\Omega, \mathbb{R})$ ($1 \leqslant p < \infty$). $W_0^{1,\infty}(\Omega, \mathbb{R}^r)$ is understood as the Sobolev space of all r-vector functions $f\colon \Omega \to \mathbb{R}^r$ with Lipschitz continuous components and zero boundary values. [16)] Jx denotes the Jacobi matrix of the vector function $x \in W_0^{1,p}(\Omega, \mathbb{R}^r)$. The diameter of a set $\Omega \subset \mathbb{R}^m$ is defined as $\mathrm{Diam}\,(\Omega) = \sup\{\, |\, x' - x''\, |\ \big|\ x', x'' \in \Omega\,\}$ while $|\,\Omega\,|$ denotes the m-dimensional Lebesgue measure of Ω. The abbreviation "$(\forall)\, t \in \mathrm{A}$" has to be read as "for almost all $s \in \mathrm{A}$" resp. "for all $s \in \mathrm{A}$ except a Lebesgue null set". Finally, the symbol $\mathfrak{o}$ denotes, depending on the context, the zero element of the underlying space.

2. Convergence of a discretization method for (P).

a) Discretization of the problem (P).

We specify the basic assumptions about the problem (P) as follows: Let $n \geqslant 1$, $m = 2$ and $1 < p < \infty$. We choose a rectangular domain $\Omega \subset \mathbb{R}^2$ with edge lengths $a, b \in \mathbb{N}$, $a \geqslant b > 0$. The integrand $f(s, \xi, v)\colon \Omega \times \mathbb{R}^n \times \mathbb{R}^{n\times 2} \to \mathbb{R}$ is measurable and essentially bounded w. r. to s and continuously differentiable w. r. to all ξ_i and v_{ij}. For all $s \in \Omega$, $B(s)$ is the (n, n)-unit matrix, and $\mathrm{G} \subset \mathbb{R}^{n\times 2}$ is the norm body

$$\mathrm{G} = \Big\{\, v \in \mathbb{R}^{n\times 2} \ \Big|\ \sum_{i=1}^{n} \sum_{j=1}^{2} |\, v_{ij}\,|^q \leqslant R^q \,\Big\} \tag{2.1}$$

with $q \in \mathbb{N}$, $q \geqslant 1$. From these assumptions, we derive immediately the existence of a feasible solution (the pair $(\mathfrak{o}, \mathfrak{o})$). For any admissible pair (x, u), the restriction $Jx(s) \in \mathrm{G}$ $(\forall)\, s \in \Omega$ implies $x \in W_0^{1,p}(\Omega, \mathbb{R}^n) \cap W_0^{1,\infty}(\Omega, \mathbb{R}^n)$. As a consequence, x possesses a Lipschitz continuous representative even in the case $1 < p \leqslant m$. If c_q denotes the constant of equivalence between the 2- and q-norm in $\mathbb{R}^2$ then all components x_i have the Lipschitz constant $C_1 = R/c_q$.

Let now $K = L = 2^N$. We construct a triangulation of Ω by first introducing a grid of $(K \times L)$ rectangles $\mathrm{Q}_{k,l}$ with edge lengths $a/2^N$ and $b/2^N$ and splitting then every rectangle along the principal diagonal into triangles $\Delta'_{k,l} = \Delta(s_{k-1,l-1}, s_{k,l-1}, s_{k,l})$ and $\Delta''_{k,l} = \Delta(s_{k-1,l-1}, s_{k,l}, s_{k-1,l})$. Thus we arrive at a regular triangulation $\mathcal{T}_N$ [17)] with vertices $s_{k,l}$ and mesh size $\sigma_N = \sqrt{a^2 + b^2}/2^N$. For every triangle, its interior angles ϑ obey $\sin\vartheta \geqslant \sin\vartheta_0 = b/\sqrt{a^2 + b^2} > 0$. [18)] The space of all piecewise affine functions with zero

---

[15)] Cf. [Droske/Rumpf 04], [Faugeras/Hermosillo 04] and [Haber/Modersitzki 07].

[16)] [Evans/Gariepy 92], p. 131, Theorem 5.

[17)] We adopt the terminology from [Goering/Roos/Tobiska 93], pp. 28 and 40, (Z1) – (Z4), and p. 138, (Z5).

[18)] Consequently, the sequence $\{\,\mathcal{T}_N\,\}$ of the triangulations satisfies the Zlámal condition, cf. [Ciarlet 87], pp. 124 and 130.

boundary values on $\partial\Omega$, which are adapted to the triangulation $\mathcal{T}_N$, will be denoted by X_0^N. Then we may restrict the problem (P) in the following way:

$$(\mathrm{P})_N\colon\quad F(x,u)=\int_\Omega f(s,x(s),u(s))\,ds \longrightarrow \mathrm{Min}\,!\,;\quad (x,u)\in\Big(W_0^{1,p}(\Omega,\mathbb{R}^n)\cap X_0^N\Big)\times L^p(\Omega,\mathbb{R}^{n\times 2})\,; \tag{2.2}$$

$$Jx(s)\;=\;u(s)\;\;(\forall)\,s\in\Omega\,; \tag{2.3}$$

$$u(s)\in \mathrm{G}\;=\;\Big\{\,v\in\mathbb{R}^{n\times 2}\;\Big|\;\sum_{i=1}^{n}\sum_{j=1}^{2}\,|\,v_{ij}\,|^q\leqslant R^q\,\Big\}\;\;(\forall)\,s\in\Omega\,. \tag{2.4}$$

For the admissible pairs (x,u) in $(\mathrm{P})_N$, it holds that

$$\int_{\mathrm{Q}_{k,l}} f(s,x(s),u(s))\,ds\;=\;\int_{\Delta'_{k,l}} f(s,x(s),u(s))\,ds\;+\;\int_{\Delta''_{k,l}} f(s,x(s),u(s))\,ds \tag{2.5}$$

$$=\int_{\Delta'_{k,l}} f\big(\,s,\,x(s),\,\begin{pmatrix}\dfrac{x_1(s_{k,l-1})-x_1(s_{k-1,l-1})}{(a/2^N)} & \dfrac{x_1(s_{k,l})-x_1(s_{k,l-1})}{(b/2^N)}\\ \vdots & \vdots\\ \dfrac{x_n(s_{k,l-1})-x_n(s_{k-1,l-1})}{(a/2^N)} & \dfrac{x_n(s_{k,l})-x_n(s_{k,l-1})}{(b/2^N)}\end{pmatrix}\big)\,ds$$

$$+\int_{\Delta''_{k,l}} f\big(\,s,\,x(s),\,\begin{pmatrix}\dfrac{x_1(s_{k,l})-x_1(s_{k-1,l})}{(a/2^N)} & \dfrac{x_1(s_{k-1,l})-x_1(s_{k-1,l-1})}{(b/2^N)}\\ \vdots & \vdots\\ \dfrac{x_n(s_{k,l})-x_n(s_{k-1,l})}{(a/2^N)} & \dfrac{x_n(s_{k-1,l})-x_n(s_{k-1,l-1})}{(b/2^N)}\end{pmatrix}\big)\,ds\,. \tag{2.6}$$

Using the abbreviation $x_i(s_{k,l})=\xi_{k,l}^{(i)}$, we obtain as the discretized problem related to $(\mathrm{P})_N$:

$$(\mathrm{D})_N\colon\quad \widetilde{F}\big(\,\xi_{1,1}^{(1)},\,\ldots\,,\xi_{K,L}^{(n)},\,v_{1,1}^{(1,1)},\,\ldots\,,v_{K,L}^{(n,4)}\,\big)\;=\;\frac{1}{2}\cdot\frac{a\,b}{4^N}$$

$$\cdot\sum_{k=1}^{K}\sum_{l=1}^{L}\Big(\,f\big(\,s_{k-1,l-1},\,\begin{pmatrix}\xi_{k-1,l-1}^{(1)}\\ \vdots\\ \xi_{k-1,l-1}^{(n)}\end{pmatrix},\,\begin{pmatrix}v_{k,l}^{(1,1)} & v_{k,l}^{(1,2)}\\ \vdots & \vdots\\ v_{k,l}^{(n,1)} & v_{k,l}^{(n,2)}\end{pmatrix}\big)\;+\;f\big(\,s_{k,l},\,\begin{pmatrix}\xi_{k,l}^{(1)}\\ \vdots\\ \xi_{k,l}^{(n)}\end{pmatrix},\,\begin{pmatrix}v_{k,l}^{(1,3)} & v_{k,l}^{(1,4)}\\ \vdots & \vdots\\ v_{k,l}^{(n,3)} & v_{k,l}^{(n,4)}\end{pmatrix}\big)\Big) \tag{2.7}$$

$$\longrightarrow \mathrm{Min}\,!\,;\quad \big(\,\xi_{0,0}^{(1)},\,\ldots\,,\xi_{K,L}^{(n)},\,v_{1,1}^{(1,1)},\,\ldots\,,v_{K,L}^{(n,4)}\,\big)\;\in\;\mathbb{R}^{n(K+1)(L+1)}\times\mathbb{R}^{4nKL}\,; \tag{2.8}$$

$$\xi_{0,l}^{(i)}\;=\;\xi_{K,l}^{(i)}\;=\;0\,,\quad 1\leqslant i\leqslant n\,,\;\;0\leqslant l\leqslant L\,; \tag{2.9}$$

$$\xi_{k,0}^{(i)}\;=\;\xi_{k,L}^{(i)}\;=\;0\,,\quad 1\leqslant i\leqslant n\,,\;\;0\leqslant k\leqslant K\,; \tag{2.10}$$

$$v_{k,l}^{(i,1)}\;=\;\frac{\xi_{k,l-1}^{(i)}-\xi_{k-1,l-1}^{(i)}}{(a/2^N)}\,;\quad v_{k,l}^{(i,2)}\;=\;\frac{\xi_{k,l}^{(i)}-\xi_{k,l-1}^{(i)}}{(b/2^N)}\,,\;\;1\leqslant i\leqslant n\,,\;\;1\leqslant k\leqslant K\,,\;\;1\leqslant l\leqslant L\,; \tag{2.11}$$

$$v_{k,l}^{(i,3)}\;=\;\frac{\xi_{k,l}^{(i)}-\xi_{k-1,l}^{(i)}}{(a/2^N)}\,;\qquad v_{k,l}^{(i,4)}\;=\;\frac{\xi_{k-1,l}^{(i)}-\xi_{k-1,l-1}^{(i)}}{(b/2^N)}\,,\;\;1\leqslant i\leqslant n\,,\;\;1\leqslant k\leqslant K\,,\;\;1\leqslant l\leqslant L\,; \tag{2.12}$$

$$\sum_{i=1}^{n}\Big(\,|\,v_{k,l}^{(i,1)}\,|^q\;+\;|\,v_{k,l}^{(i,2)}\,|^q\Big)\;\leqslant\;R^q\,,\;\;1\leqslant k\leqslant K\,,\;\;1\leqslant l\leqslant L\,; \tag{2.13}$$

$$\sum_{i=1}^{n}\Big(\,|\,v_{k,l}^{(i,3)}\,|^q\;+\;|\,v_{k,l}^{(i,4)}\,|^q\Big)\;\leqslant\;R^q\,,\;\;1\leqslant k\leqslant K\,,\;\;1\leqslant l\leqslant L\,. \tag{2.14}$$

In consequence of the possible discontinuity of the control variables within $(\mathrm{P})_N$, the number of the corresponding discretization variables in $(\mathrm{D})_N$ has been doubled. For the same reason, the classical two-dimensional Newton-Cotes cubature formula [19] had to be modified.

[19] Cf. [Maess 88], p. 238 f.

b) Existence of global minimizers for the problems (P)**,** $(\mathrm{P})_N$ **and** $(\mathrm{D})_N$**.**

Theorem 2.1. (Global minimizers for (P) **and** $(\mathrm{P})_N$**)** *Consider the problem* (P) *under the assumptions of Section 2.a). Assume further that the integrand* $f(s,\xi,v)$ *is convex as a function of* v *for almost all* $s \in \Omega$ *and all* $\xi \in \mathbb{R}^n$*, and a growth condition*

$$\big|\, f(s,\xi,v) \,\big| \leqslant \varphi_1(s) + \varphi_2(\,|\,\xi\,|\,,\,|\,v\,|\,) \quad (\forall)\, s \in \Omega \quad \forall\,(\xi,v) \in \mathbb{R}^n \times \mathrm{G} \tag{2.15}$$

with $\varphi_1 \in L^1(\Omega,\mathbb{R})$*,* $\varphi_1(s) \geqslant 0$ $(\forall)\, s \in \Omega$ *and* $\varphi_2 \in C^0(\mathbb{R}^n \times \mathrm{G}\,,\,\mathbb{R})$*,* $\varphi_2(\,|\,\xi\,|\,,\,|\,v\,|\,) \geqslant 0$ $\forall\,(\xi,v) \in \mathbb{R}^n \times \mathrm{G}$ *is satisfied where* φ_2 *is a monotonically increasing function in* $|\,\xi\,|$ *as well as in* $|\,v\,|$*.*

1)[20] *Then* (P) *admits a global minimizer* $(\hat{x},\hat{u})$*.*

2) Every problem $(\mathrm{P})_N$*,* $N \in \mathbb{N}$*, admits a global minimizer* $(\hat{x}_N,\hat{u}_N)$*.*

Proof. 2) Although the feasible domain is restricted to $\big(W_0^{1,p}(\Omega,\mathbb{R}^n) \cap X_0^N\big) \times L^p(\Omega,\mathbb{R}^{n\times 2})$, the arguments from [Pickenhain/Wagner 00], pp. 222 – 224, and [Dacorogna 08], p. 378, Theorem 8.8., remain valid. ∎

Theorem 2.2. *Under the assumptions of Theorem 2.1., every problem* $(\mathrm{D})_N$*,* $N \in \mathbb{N}$*, admits a global minimizer* $\big(\hat{\xi}_{0,0}^{(1)},\ \ldots,\ \hat{\xi}_{K,L}^{(n)},\ \hat{v}_{1,1}^{(1,1)},\ \ldots,\ \hat{v}_{K,L}^{(n,4)}\big)$*.*

Proof. The feasible domain of $(\mathrm{D})_N$ is nonempty, convex and compact since the equations (2.11) − (2.12) for $v_{k,l}^{(i,j)}$ and the assignment $\xi_{k,0}^{(i)} = \xi_{k,L}^{(i)} = \xi_{0,l}^{(i)} = \xi_{K,l}^{(i)} = 0$ at the boundary vertices imply the boundedness of the variables $\xi_{k,l}^{(i)}$ as well. The objective $\widetilde{F}$ is continuous w. r. to all variables; consequently, it takes its global minimum on the feasible domain. ∎

c) A convergence theorem for the discretization method.

Theorem 2.3. (Convergence of the discretization method for (P)**)** *Consider the problems* (P)*,* $(\mathrm{P})_N$ *and* $(\mathrm{D})_N$ *under the assumptions of Theorem 2.1. Let* $1 < p < \infty$ *and assume further that* $4\,C_1\,\sigma_N/\sin\vartheta_0 \leqslant 1$ *as well as* $8\,n\,C_1\,(1+C_1)^{q-1}\,\sigma_N/\sin\vartheta_0 < 1$ *with* $C_1 = R/c_q$*.*

1) **(Relations between** (P) **and** $(\mathrm{P})_N$**)** *Let* $N \in \mathbb{N}$*. For all global minimizers* $(\hat{x},\hat{u})$ *of* (P) *and* $(\hat{x}_N,\hat{u}_N)$ *of* $(\mathrm{P})_N$*, the inequality*

$$F(\hat{x},\hat{u}) \leqslant F(\hat{x}_N,\hat{u}_N) \leqslant F(\hat{x},\hat{u}) + C_2\,\sigma_N \tag{2.16}$$

holds with a constant $C_2 > 0$*, which does not depend on* $N \in \mathbb{N}$*. Moreover, every sequence* $\{\,(\hat{x}_N,\hat{u}_N)\,\}$ *of global minimizers of* $(\mathrm{P})_N$ *contains a subsequence* $\{\,(\hat{x}_{N'},\hat{u}_{N'})\,\}$ *with*

$$\big\|\,\hat{x}_{N'} - \hat{x}\,\big\|_{C^0(\Omega,\mathbb{R}^n)} \to 0 \quad \textit{and} \quad \big(\,\hat{u}_{N'} - \hat{u}\,\big) \rightharpoonup \mathfrak{o}_{L^p(\Omega,\mathbb{R}^{n\times 2})} \tag{2.17}$$

where $(\hat{x},\hat{u})$ *is a global minimizer of* (P)*, and the entire sequence satisfies*

$$\big|\, F(\hat{x}_N,\hat{u}_N) - F(\hat{x},\hat{u}) \,\big| \to 0\,. \tag{2.18}$$

2) **(Relations between** (P) **and** $(\mathrm{D})_N$**)** *Let* $N \in \mathbb{N}$*. Assume that for all feasible solutions* (x,u) *of* $(\mathrm{P})_N$ *and all indices* $1 \leqslant k \leqslant K$*,* $1 \leqslant l \leqslant L$ *the conditions*

$$\big|\, f(s,x(s),u(s)) - f(s_{k-1,l-1},x(s),u(s)) \,\big| \leqslant C_3 \cdot \sigma_N \quad (\forall)\, s \in \Delta'_{k,l} \quad \textit{and} \tag{2.19}$$

$$\big|\, f(s,x(s),u(s)) - f(s_{k,l},x(s),u(s)) \,\big| \leqslant C_3 \cdot \sigma_N \quad (\forall)\, s \in \Delta''_{k,l} \tag{2.20}$$

[20] [Brune/Maurer/Wagner 09], p. 1195, Theorem 3.1.

hold. Consider a global minimizer $(\hat{\xi}, \hat{v})$ *of* $(\mathrm{D})_N$ *and the related feasible solution* $(\widetilde{x}_N, \widetilde{u}_N)$ *of* $(\mathrm{P})_N$ *defined by*

$$\widetilde{x}_{N,i}(s_{k,l}) \,=\, \hat{\xi}^{(i)}_{k,l}\,, \quad 1 \leqslant i \leqslant n\,,\ 0 \leqslant k \leqslant K\,,\ 0 \leqslant l \leqslant L\,. \tag{2.21}$$

Then for all global minimizers $(\hat{x}, \hat{u})$ *of* (P)*, the inequality*

$$F(\hat{x}, \hat{u}) \,\leqslant\, F(\widetilde{x}_N, \widetilde{u}_N) \,\leqslant\, F(\hat{x}, \hat{u}) \,+\, C_4\, \sigma_N \tag{2.22}$$

holds with a constant $C_4 > 0$*, which does not depend on* $N \in \mathbb{N}$*. Moreover, the sequence* $\{\, (\widetilde{x}_N, \widetilde{u}_N)\,\}$ *contains a subsequence* $\{\, (\widetilde{x}_{N'}, \widetilde{u}_{N'})\,\}$ *with*

$$\big\|\, \widetilde{x}_{N'} - \hat{x}\,\big\|_{C^0(\Omega,\mathbb{R}^n)} \to 0 \quad \textit{and} \quad \big(\, \widetilde{u}_{N'} - \hat{u}\,\big) \rightharpoonup \mathfrak{o}_{L^p(\Omega,\mathbb{R}^{n\times 2})} \tag{2.23}$$

where $(\hat{x}, \hat{u})$ *is a global minimizer of* (P)*, and the entire sequence satisfies*

$$\big|\, F(\widetilde{x}_N, \widetilde{u}_N) - F(\hat{x}, \hat{u})\,\big| \to 0\,. \tag{2.24}$$

Remarks. 1. The additional assumption in Part 2) of the theorem reflects the fact that the integrand f is only a measurable and essentially bounded function of s.

2. If (P) admits an uniquely determined global minimizer then the convergence relations (2.17) and (2.23) hold for the entire sequences $\{\, (\hat{x}_N, \hat{u}_N)\,\}$ resp. $\{\, (\widetilde{x}_N, \widetilde{u}_N)\,\}$, since, in consequence of (2.18) and (2.24), all weakly convergent subsequences possess the same limit element. [21)]

3. Both inequalities for C_1, σ_N and ϑ_0 are required in order to ensure the relations (2.33) and (2.34) below.

Proof. The proof will be divided into seven steps.

• **Step 1.** C^∞*-approximation of a global minimizer of* (P). From Theorem 2.1., we know that (P) possesses a global minimizer $(\hat{x}, \hat{u}) \in W^{1,\infty}_0(\Omega, \mathbb{R}^n) \times L^\infty(\Omega, \mathbb{R}^{n\times 2})$. We rely to the following approximation theorem:

Theorem 2.4. [22)] *Assume that* $\Omega \subset \mathbb{R}^m$ *is the closure of a bounded strongly Lipschitz domain,* $\mathrm{G} \subset \mathbb{R}^{nm}$ *is a convex, compact set with* $\mathfrak{o} \in \mathrm{int}\,(\mathrm{G})$ *and* $\hat{x} \in W^{1,\infty}_0(\Omega, \mathbb{R}^n)$ *is a function with* $J\hat{x}(s) \in \mathrm{G}$ *for almost all* $s \in \Omega$*. Then* $\hat{x}$ *can be approximated by a sequence of functions* $z^N \in C^\infty_0(\Omega, \mathbb{R}^n)$ *with the following properties:*

1) $\displaystyle\lim_{N\to\infty} \big\|\, z^N - \hat{x}\,\big\|_{C^0(\Omega,\mathbb{R}^n)} \,=\, 0\,,$ (2.25)

2) $\displaystyle\lim_{N\to\infty} \big\|\, Jz^N - J\hat{x}\,\big\|_{L^1(\Omega,\mathbb{R}^{nm})} \,=\, 0\,,$ (2.26)

3) $Jz^N(s) \in \mathrm{G}$ *for all* $s \in \Omega$. (2.27)

Consequently, in relation to $\hat{x}$ we may choose a function $z \in C^\infty_0(\Omega, \mathbb{R}^{n\times 2})$ with

$$\big\|\, z - \hat{x}\,\big\|_{C^0(\Omega,\mathbb{R}^n)} \,+\, \big\|\, Jz - J\hat{x}\,\big\|_{L^1(\Omega,\mathbb{R}^{n\times 2})} \,\leqslant\, \sigma_N \tag{2.28}$$

and $Jz(s) \in \mathrm{G}$ for all $s \in \Omega$. The Lipschitz constant of the components of z amounts to $C_1 = R/c_q$ as well, and the zero boundary conditions for $\hat{x}$ and z imply

$$\big\|\, \hat{x}_i\,\big\|_{C^0(\Omega,\mathbb{R})} \,\leqslant\, C_1 \cdot \mathrm{Diam}\,(\Omega) \,=\, \frac{C_1}{2}\,\sqrt{a^2 + b^2} \quad \text{and} \quad \big\|\, z_i\,\big\|_{C^0(\Omega,\mathbb{R})} \,\leqslant\, C_1 \cdot \mathrm{Diam}\,(\Omega)\,,\ 1 \leqslant i \leqslant n\,. \tag{2.29}$$

[21)] [Gajewski/Gröger/Zacharias 74], p. 10, Lemma 5.4.

[22)] [Wagner 99], p. 2, Theorem 1.5., with $\mathrm{S}(s) \equiv \mathrm{G}$, $\Gamma = \partial\Omega$ and $c = \mathfrak{o}$.

• **Step 2.** *Piecewise affine approximation of z.* We approximate z by that continuous, piecewise affine function $y \in X_0^N$, which coincides with z in all vertices of the the triangulation $\mathcal{T}_N$. Let us first note that

$$\| y_i \|_{C^0(\Omega,\mathbb{R})} \leqslant \| z_i \|_{C^0(\Omega,\mathbb{R})} \leqslant C_1 \cdot \mathrm{Diam}\,(\Omega)\,,\ 1 \leqslant i \leqslant n\,. \tag{2.30}$$

Moreover, the following inequalities hold:

Lemma 2.5. [23] *Let a function $z \in C_0^\infty(\Omega, \mathbb{R}^{n\times 2})$ with Lipschitz constant $C_1 > 0$ for all components be given. Then for $1 \leqslant i \leqslant n$, the following estimates hold:*

$$\Big| z_i(s) - y_i(s) \Big| \leqslant C_1 \cdot C_5 \cdot \sigma_N \quad \textit{with} \quad C_5 = 6 + \frac{32}{\sin\vartheta_0}\,; \tag{2.31}$$

$$\Big| \frac{\partial z_i}{\partial s_j}(s) - \frac{\partial y_i}{\partial s_j}(s) \Big| \leqslant C_1 \cdot C_6 \cdot \sigma_N \quad \textit{with} \quad C_6 = \frac{4}{\sin\vartheta_0}\,. \tag{2.32}$$

From Lemma 2.5. we conclude that, for all sufficiently large $N \in \mathbb{N}$, every pair $\big((1-D)\,y\,,\ (1-D)\,Jy \big)$ with

$$0 < D = 2\,n\,C_1\,(1+C_1)^{q-1}\,C_6\,\sigma_N < 1 \tag{2.33}$$

is admissible in $(\mathrm{P})_N$. To see this, we enlarge N until

$$C_1 \cdot C_6 \cdot \sigma_N \leqslant 1 \tag{2.34}$$

is satisfied. Then it follows that

$$\begin{aligned}
\Big| \frac{\partial y_i}{\partial s_j}(s) \Big|^q &\leqslant \Big(\Big| \frac{\partial z_i}{\partial s_j}(s) \Big| + \Big| \frac{\partial y_i}{\partial s_j}(s) - \frac{\partial z_i}{\partial s_j}(s) \Big| \Big)^q \\
&= \Big| \frac{\partial z_i}{\partial s_j}(s) \Big|^q + \sum_{k=1}^{q-1} \binom{q}{k} \cdot \Big| \frac{\partial z_i}{\partial s_j}(s) \Big|^k \cdot \Big| \frac{\partial y_i}{\partial s_j}(s) - \frac{\partial z_i}{\partial s_j}(s) \Big|^{q-k} && (2.35)\\
&\leqslant \Big| \frac{\partial z_i}{\partial s_j}(s) \Big|^q + \Big| \frac{\partial y_i}{\partial s_j}(s) - \frac{\partial z_i}{\partial s_j}(s) \Big| \cdot \sum_{k=1}^{q-1} \binom{q}{k} \cdot \Big| \frac{\partial z_i}{\partial s_j}(s) \Big|^k \cdot 1^{q-1-k} && (2.36)\\
&\leqslant \Big| \frac{\partial z_i}{\partial s_j}(s) \Big|^q + \Big| \frac{\partial y_i}{\partial s_j}(s) - \frac{\partial z_i}{\partial s_j}(s) \Big| \cdot \sum_{k=1}^{q-1} \binom{q}{k} \cdot L^k \cdot 1^{q-k} && (2.37)\\
&\leqslant \Big| \frac{\partial z_i}{\partial s_j}(s) \Big|^q + C_1\,(1+C_1)^{q-1}\,C_6\,\sigma_N \quad \Longrightarrow && (2.38)\\
\sum_{i=1}^{n} \sum_{j=1}^{2} \Big| \frac{\partial y_i}{\partial s_j}(s) \Big|^q &\leqslant R^q + 2\,n\,C_1\,(1+C_1)^{q-1}\,C_6\,\sigma_N \quad \Longrightarrow && (2.39)\\
(1-D) \sum_{i=1}^{n} \sum_{j=1}^{2} \Big| \frac{\partial y_i}{\partial s_j}(s) \Big|^q &\leqslant (1-D)\,R^q + 2\,n\,C_1\,(1+C_1)^{q-1}\,C_6\,\sigma_N\,. && (2.40)
\end{aligned}$$

From the requirement

$$(1-D)\,R^q + 2\,n\,C_1\,(1+C_1)^{q-1}\,C_6\,\sigma_N \leqslant R^q\,, \tag{2.41}$$

we obtain

$$D = 2\,n\,C_1\,(1+C_1)^{q-1}\,C_6\,\sigma_N \tag{2.42}$$

[23] [Wagner 03], p. 41, Lemma 0.1. and 0.2., modified from [Ekeland/Témam 99], p. 309, Proposition 2.1.

as well as

$$\Big|\, z_i(s) - (1-D)\, y_i(s) \,\Big| \;\leqslant\; C_1 \cdot C_5 \cdot \sigma_N \;+\; D \cdot \Big|\, y_i(s) \,\Big|$$
$$\leqslant \Big(C_1\, C_5 \;+\; 2\, n\, C_1\, (1+C_1)^{q-1}\, C_6 \cdot C_1 \cdot \mathrm{Diam}\,(\Omega) \Big) \cdot \sigma_N \;=\; C_7\, \sigma_N\,; \quad (2.43)$$

$$\Big|\, \frac{\partial z_i}{\partial s_j}(s) - (1-D)\, \frac{\partial y_i}{\partial s_j}(s) \,\Big| \;\leqslant\; C_1 \cdot C_6 \cdot \sigma_N \;+\; D \,\Big|\, \frac{\partial y_i}{\partial s_j}(s) \,\Big|$$
$$\leqslant \Big(C_1\, C_6 \;+\; 2\, n\, C_1\, (1+C_1)^{q-1}\, C_6 \cdot R \Big) \cdot \sigma_N \;=\; C_8\, \sigma_N\,. \quad (2.44)$$

Now we define $\widetilde{y} = (1-D)\, y$ and $\widetilde{w} = (1-D)\, Jy$.

• **Step 3.** *Estimation of the objective values in* (P) *and* $(\mathrm{P})_N$. From the assumed differentiability of the integrand $f(s,\xi,v)$ w. r. to ξ_i and v_{ij}, we obtain together with (2.29):

$$\big|\, f(s,\xi',v) \;-\; f(s,\xi'',v) \,\big| \;\leqslant\; \Big(\max_{(s,\,\xi,\,v)\,\in\,\Omega\,\times\,\mathrm{A}\,\times\,\mathrm{G}} \big|\, \nabla_\xi\, f(s,\xi,v) \,\big| \Big) \cdot \big|\, \xi' - \xi'' \,\big|$$
$$= C_9 \cdot \big|\, \xi' - \xi'' \,\big| \quad (\forall)\, s \in \Omega \;\; \forall\, \xi',\, \xi'' \in \mathrm{A} \;\; \forall\, v \in \mathrm{G}\,; \quad (2.45)$$

$$\big|\, f(s,\xi,v') \;-\; f(s,\xi,v'') \,\big| \;\leqslant\; \Big(\max_{(s,\,\xi,\,v)\,\in\,\Omega\,\times\,\mathrm{A}\,\times\,\mathrm{G}} \big|\, \nabla_v\, f(s,\xi,v) \,\big| \Big) \cdot \big|\, v' - v'' \,\big|$$
$$= C_{10} \cdot \big|\, v' - v'' \,\big| \quad (\forall)\, s \in \Omega \;\; \forall\, \xi \in \mathrm{A} \;\; \forall\, v',\, v'' \in \mathrm{G}\,. \quad (2.46)$$

Here $\mathrm{A} \subset \mathbb{R}^n$ denotes the closed ball with center $\mathfrak{o}$ and radius $(\, C_1 \cdot \mathrm{Diam}\,(\Omega)\,)$ while the vector $\nabla_v\, f(s,\xi,v)$ has been assembled from the columns of the Jacobi matrix with entries $\partial f(s,\xi,v)/\partial v_{ij}$. It follows that

$$F(\hat{x},\hat{u}) \;\leqslant\; F(\hat{x}_N,\hat{u}_N) \;\leqslant\; F(\widetilde{y},\widetilde{w}) \;\leqslant\; F(\hat{x},\hat{u}) \;+\; \big|\, F(z,Jz) - F(\hat{x},\hat{u}) \,\big| \;+\; \big|\, F(\widetilde{y},\widetilde{w}) - F(z,Jz) \,\big| \quad (2.47)$$

$$\leqslant\; F(\hat{x},\hat{u}) \;+\; \int_\Omega \Big(\big|\, f(s,z(s),Jz(s)) - f(s,\hat{x}(s),Jz(s)) \,\big| \;+\; \big|\, f(s,\hat{x}(s),Jz(s)) - f(s,\hat{x}(s),\hat{u}(s)) \,\big| \Big)\, ds$$
$$+\; \int_\Omega \Big(\big|\, f(s,\widetilde{y}(s),\widetilde{w}(s)) - f(s,z(s),\widetilde{w}(s)) \,\big| \;+\; \big|\, f(s,z(s),\widetilde{w}(s)) - f(s,z(s),Jz(s)) \,\big| \Big)\, ds \quad (2.48)$$

$$\leqslant\; F(\hat{x},\hat{u}) \;+\; \int_\Omega \Big(C_9\, \big|\, z(s) - \hat{x}(s) \,\big| \;+\; C_{10}\, \big|\, Jz(s) - \hat{u}(s) \,\big| \Big)\, ds$$
$$+\; \int_\Omega \Big(C_9\, \big|\, \widetilde{y}(s) - z(s) \,\big| \;+\; C_{10}\, \big|\, \widetilde{w}(s) - Jz(s) \,\big| \Big)\, ds \quad (2.49)$$

$$\leqslant\; F(\hat{x},\hat{u}) \;+\; \Big(|\,\Omega\,|\, C_9\, \sqrt{n} \;+\; C_{10}\, \sqrt{2\,n} \;+\; |\,\Omega\,|\, C_7\, C_9\, \sqrt{n} \;+\; |\,\Omega\,|\, C_8\, C_{10}\, \sqrt{2\,n} \Big) \cdot \sigma_N\,. \quad (2.50)$$

The first and the second member have been estimated by use of (2.28); the third and the fourth one by use of (2.43) and (2.44).

• **Step 4.** *Analysis of the sequence* $\{\, (\hat{x}_N,\hat{u}_N)\,\}$. For every $N \in \mathbb{N}$, we choose a global minimizer $(\hat{x}_N,\hat{u}_N)$ of $(\mathrm{P})_N$. All pairs $(\hat{x}_N,\hat{u}_N)$ are feasible in (P) as well, and by Step 3, the sequence $\{\, (\hat{x}_N,\hat{u}_N)\,\}$, $W_0^{1,p}(\Omega,\mathbb{R}^n) \times L^p(\Omega,\mathbb{R}^{n\times 2})$ is a minimizing sequence for (P). Due to (1.3) and (2.29), the feasible domain of (P) is bounded. Consequently, $\{\, (\hat{x}_N,\hat{u}_N)\,\}$ admits a weakly convergent subsequence whose limit $(\hat{x},\hat{u})$ is feasible in (P) as well (this holds in analogy to [Wagner 96], p. 60, Lemma 4.2–10). Since $\hat{x}_N \in W_0^{1,p}(\Omega,\mathbb{R}^n) \cap W_0^{1,\infty}(\Omega,\mathbb{R}^n)$ for all $N \in \mathbb{N}$, we may assume that $p > 2$, and from the Rellich-Kondrachev embedding theorem [24] we get the existence of a subsequence with uniform convergence w. r. to x. The proof of Part 1) is complete.

[24] [Adams/Fournier 07], p. 168, Theorem 6.3.

• **Step 5.** *Relations between feasible solutions of* $(\mathrm{P})_N$ *and* $(\mathrm{D})_N$. The proof of Part 2) starts with the following

Lemma 2.6. *Consider the problems* $(\mathrm{P})_N$ *and* $(\mathrm{D})_N$ *under the assumptions of Theorem 2.1. Assume further that the conditions*

$$\big|\, f(s,x(s),u(s)) - f(s_{k-1,l-1},x(s),u(s))\,\big| \;\leqslant\; C_3\cdot\sigma_N \;\; (\forall)\, s\in\Delta'_{k,l} \quad \textit{and} \tag{2.51}$$

$$\big|\, f(s,x(s),u(s)) - f(s_{k,l},x(s),u(s))\,\big| \;\leqslant\; C_3\cdot\sigma_N \;\; (\forall)\, s\in\Delta''_{k,l} \tag{2.52}$$

hold for all feasible solutions (x,u) *of* $(\mathrm{P})_N$ *and all indices* $1\leqslant k\leqslant K$, $1\leqslant l\leqslant L$. *Then for every feasible solution* (x,u) *of* $(\mathrm{P})_N$ *there exists a feasible solution* $\big(\,\xi^{(1)}_{0,0},\;\ldots,\,\xi^{(n)}_{K,L},\,v^{(1,1)}_{1,1},\;\ldots,\,v^{(n,4)}_{K,L}\,\big)$ *of* $(\mathrm{D})_N$ *with*

$$\big|\, F(x,u) - \widetilde{F}\big(\,\xi^{(1)}_{0,0},\;\ldots,\,\xi^{(n)}_{K,L},\,v^{(1,1)}_{1,1},\;\ldots,\,v^{(n,4)}_{K,L}\,\big)\,\big| \;\leqslant\; C_{11}\cdot\sigma_N\,. \tag{2.53}$$

Conversely, for every feasible solution $\big(\,\xi^{(1)}_{0,0},\;\ldots,\,\xi^{(n)}_{K,L},\,v^{(1,1)}_{1,1},\;\ldots,\,v^{(n,4)}_{K,L}\,\big)$ *of* $(\mathrm{D})_N$ *there exists a feasible solution* (x,u) *of* $(\mathrm{P})_N$ *with*

$$\big|\, \widetilde{F}\big(\,\xi^{(1)}_{0,0},\;\ldots,\,\xi^{(n)}_{K,L},\,v^{(1,1)}_{1,1},\;\ldots,\,v^{(n,4)}_{K,L}\,\big) - F(x,u)\,\big| \;\leqslant\; C_{11}\cdot\sigma_N\,. \tag{2.54}$$

Proof. Let a feasible solution (x,u) of $(\mathrm{P})_N$ be given. The objective in $(\mathrm{P})_N$ may be represented as

$$F(x,u) \;=\; \sum_{k=1}^{K}\sum_{l=1}^{L}\Big(\int_{\Delta'_{k,l}} f(s,x(s),u(s))\,ds \;+\; \int_{\Delta''_{k,l}} f(s,x(s),u(s))\,ds\Big)\,. \tag{2.55}$$

With the setting

$$\xi^{(i)}_{k,l} \;=\; x_i(s_{k,l})\,,\;\; 1\leqslant i\leqslant n\,,\;\; 0\leqslant k\leqslant K\,,\;\; 0\leqslant l\leqslant L\,; \tag{2.56}$$

$$v^{(i,1)}_{k,l} \;=\; \frac{x_i(s_{k,l-1})-x_i(s_{k-1,l-1})}{(a/2^N)}\,;\;\; v^{(i,2)}_{k,l} \;=\; \frac{x_i(s_{k,l})-x_i(s_{k,l-1})}{(b/2^N)}\,,\;\; 1\leqslant i\leqslant n\,,\;\; 1\leqslant k\leqslant K\,,\;\; 1\leqslant l\leqslant L\,;$$

$$v^{(i,3)}_{k,l} \;=\; \frac{x_i(s_{k,l})-x_i(s_{k-1,l})}{(a/2^N)}\,;\qquad v^{(i,4)}_{k,l} \;=\; \frac{x_i(s_{k-1,l}-x_i(s_{k-1,l-1})}{(b/2^N)}\,,\;\; 1\leqslant i\leqslant n\,,\;\; 1\leqslant k\leqslant K\,,\;\; 1\leqslant l\leqslant L\,,$$

we obtain a feasible solution $\big(\,\xi^{(1)}_{0,0},\;\ldots,\,\xi^{(n)}_{K,L},\,v^{(1,1)}_{1,1},\;\ldots,\,v^{(n,4)}_{K,L}\,\big)$ of $(\mathrm{D})_N$, and the difference of the objectives can be estimated as follows:

$$\begin{aligned}
&\Big|\, F(x,u) - \widetilde{F}\big(\,\xi^{(1)}_{0,0},\;\ldots,\,\xi^{(n)}_{K,L},\,v^{(1,1)}_{1,1},\ldots,\,v^{(n,4)}_{K,L}\,\big)\,\Big| \\
&\leqslant \sum_{k=1}^{K}\sum_{l=1}^{L}\Big(\int_{\Delta'_{kl}}\Big|\, f\big(\,s,\,x(s),\,\begin{pmatrix} v^{(1,1)}_{k,l} & v^{(1,2)}_{k,l}\\ \vdots & \vdots\\ v^{(n,1)}_{k,l} & v^{(n,2)}_{k,l}\end{pmatrix}\big) \;-\; f\big(\,s_{k-1,l-1},\,x(s),\,\begin{pmatrix} v^{(1,1)}_{k,l} & v^{(1,2)}_{k,l}\\ \vdots & \vdots\\ v^{(n,1)}_{k,l} & v^{(n,2)}_{k,l}\end{pmatrix}\big)\,\Big|\,ds \\
&\quad + \int_{\Delta'_{kl}}\Big|\, f\big(\,s_{k-1,l-1},\,\begin{pmatrix} x_1(s)\\ \vdots\\ x_n(s)\end{pmatrix},\,\begin{pmatrix} v^{(1,1)}_{k,l} & v^{(1,2)}_{k,l}\\ \vdots & \vdots\\ v^{(n,1)}_{k,l} & v^{(n,2)}_{k,l}\end{pmatrix}\big) \;-\; f\big(\,s_{k-1,l-1},\,\begin{pmatrix} \xi^{(1)}_{k-1,l-1}\\ \vdots\\ \xi^{(n)}_{k-1,l-1}\end{pmatrix},\,\begin{pmatrix} v^{(1,1)}_{k,l} & v^{(1,2)}_{k,l}\\ \vdots & \vdots\\ v^{(n,1)}_{k,l} & v^{(n,2)}_{k,l}\end{pmatrix}\big)\,\Big|\,ds
\end{aligned}$$

$$+\int_{\Delta''_{kl}}\Big|\,f\big(\,s,\,x(s),\,\begin{pmatrix} v_{k,l}^{(1,3)} & v_{k,l}^{(1,4)} \\ \vdots & \vdots \\ v_{k,l}^{(n,3)} & v_{k,l}^{(n,4)} \end{pmatrix}\big)\,-\,f\big(\,s_{k,l},\,x(s),\,\begin{pmatrix} v_{k,l}^{(1,3)} & v_{k,l}^{(1,4)} \\ \vdots & \vdots \\ v_{k,l}^{(n,3)} & v_{k,l}^{(n,4)} \end{pmatrix}\big)\,\Big|\,ds$$

$$+\int_{\Delta''_{kl}}\Big|\,f\big(\,s_{k,l},\,\begin{pmatrix} x_1(s) \\ \vdots \\ x_n(s) \end{pmatrix},\,\begin{pmatrix} v_{k,l}^{(1,3)} & v_{k,l}^{(1,4)} \\ \vdots & \vdots \\ v_{k,l}^{(n,3)} & v_{k,l}^{(n,4)} \end{pmatrix}\big)\,-\,f\big(\,s_{k,l},\,\begin{pmatrix} \xi_{k,l}^{(1)} \\ \vdots \\ \xi_{k,l}^{(n)} \end{pmatrix},\,\begin{pmatrix} v_{k,l}^{(1,3)} & v_{k,l}^{(1,4)} \\ \vdots & \vdots \\ v_{k,l}^{(n,3)} & v_{k,l}^{(n,4)} \end{pmatrix}\big)\,\Big|\,ds\Big) \tag{2.57}$$

$$\leqslant \sum_{k=1}^{K}\sum_{l=1}^{L}\Big(\,\big|\,\Delta'_{kl}\,\big|\cdot C_3\cdot\sigma_N\,+\,\big|\,\Delta'_{kl}\,\big|\cdot C_9\cdot\Big|\begin{pmatrix} x_1(s)-\xi_{k-1,l-1}^{(1)} \\ \vdots \\ x_n(s)-\xi_{k-1,l-1}^{(n)} \end{pmatrix}\Big|$$

$$+\,\big|\,\Delta''_{kl}\,\big|\cdot C_3\cdot\sigma_N\,+\,\big|\,\Delta''_{kl}\,\big|\cdot C_9\cdot\Big|\begin{pmatrix} x_1(s)-\xi_{k,l}^{(1)} \\ \vdots \\ x_n(s)-\xi_{k,l}^{(n)} \end{pmatrix}\Big|\,\Big) \tag{2.58}$$

$$\leqslant\,\Big(\,\big|\,\Omega\,\big|\cdot C_3\,+\,\big|\,\Omega\,\big|\cdot C_9\cdot C_1\cdot\sqrt{n}\cdot 2\,\Big)\cdot\sigma_N\,=\,C_{11}\,\sigma_N\,. \tag{2.59}$$

Conversely, to every feasible solution $\big(\,\xi_{0,0}^{(1)},\,\ldots,\,\xi_{K,L}^{(n)},\,v_{1,1}^{(1,1)},\,\ldots,\,v_{K,L}^{(n,4)}\,\big)$ of $(\mathrm{D})_N$, we may relate immediately a feasible solution (x,u) of $(\mathrm{P})_N$ defined by $x_i(s_{kl})=\xi_{kl}^{(i)}$, $1\leqslant i\leqslant n$, $0\leqslant k\leqslant K$, $0\leqslant l\leqslant L$. Then for the objective values, the same estimate as above holds. ∎

• **Step 6.** *Estimation of the objective values of* (P) *and* $(\mathrm{D})_N$. We consider a global minimizer $\big(\,\hat\xi_{0,0}^{(1)},\,\ldots,\,\hat\xi_{K,L}^{(n)},\,\hat v_{1,1}^{(1,1)},\,\ldots,\,\hat v_{K,L}^{(n,4)}\,\big)$ of $(\mathrm{D})_N$ and the related feasible solution $(\widetilde x_N,\widetilde u_N)$ of $(\mathrm{P})_N$ defined by $\widetilde x_{N,_i}(s_{k,l})=\hat\xi_{k,l}^{(i)}$ and $\widetilde u_{N,ij}(s)=\partial\widetilde x_{N,i}(s)/\partial s_j$. Further, let $(\hat x,\hat u)$ be a global minimizer of (P) and $(\hat x_N,\hat u_N)$ a global minimizer of $(\mathrm{P})_N$. By (ξ,v) we denote those feasible solution of $(\mathrm{D})_N$, which is defined in analogy to (2.56) by $\xi_{k,l}^{(i)}=\hat x_{N,i}(s_{k,l})$. Then from Part 1) and Lemma 2.6. it follows that

$$\begin{aligned} F(\hat x,\hat u)\,&\leqslant\,F(\hat x_N,\hat u_N)\,\leqslant\,F(\widetilde x_N,\widetilde u_N)\,\leqslant\,\widetilde F(\hat\xi,\hat v)\,+\,C_{11}\,\sigma_N\,\leqslant\,\widetilde F(\xi,v)\,+\,C_{11}\,\sigma_N \\ &\leqslant\,F(\hat x_N,\hat u_N)\,+\,2\,C_{11}\,\sigma_N\,\leqslant\,F(\hat x,\hat u)\,+\,\big(\,C_2+2\,C_{11}\,\big)\cdot\sigma_N\,=\,F(\hat x,\hat u)\,+\,C_4\,\sigma_N\,. \end{aligned} \tag{2.60}$$

• **Step 7.** *Analysis of the sequence* $\{\,(\widetilde x_n,\widetilde u_N)\,\}$. All pairs $(\widetilde x_N,\widetilde u_N)$ are feasible in (P), and from (2.60) we may conclude that $\{\,(\widetilde x_N,\widetilde u_N)\,\}$, $W_0^{1,p}(\Omega,\mathbb{R}^n)\,\times\,L^p(\Omega,\mathbb{R}^{n\times 2})$ forms a minimizing sequence in (P) as well. The existence of a subsequence with the claimed properties can be ensured as in Step 4, and the proof of Part 2) is complete. ∎

Corollary 2.7. (Maximum norm restrictions for the control) *Let* $1<p<\infty$ *and consider again the problems* (P), $(\mathrm{P})_N$ *and* $(\mathrm{D})_N$ *under the assumptions of Theorem 2.1. If the norm body* (2.1) *in* (P) *and* $(\mathrm{P})_N$ *is replaced by*

$$\mathrm{G}\,=\,\Big\{\,v\in\mathbb{R}^{n\times 2}\;\Big|\;\operatorname*{Max}_{1\leqslant i\leqslant n}\;\operatorname*{Max}_{j\,=\,1,\,2}\,|\,v_{ij}\,|\,\leqslant\,R\,\Big\} \tag{2.61}$$

and the restrictions (2.13) *and* (2.14) *in* $(\mathrm{D})_N$ *are replaced accordingly by*

$$\operatorname*{Max}_{1\leqslant i\leqslant n}\Big(\,|\,v_{k,l}^{(i,1)}\,|\,,\,|\,v_{k,l}^{(i,2)}\,|\,,\,|\,v_{k,l}^{(i,3)}\,|\,,\,|\,v_{k,l}^{(i,4)}\,|\,\Big)\,\leqslant R \tag{2.62}$$

then Theorems 2.1., 2.2. and 2.3. remain true for the modified problems (in Theorem 2.3., one has to choose now $C_1=\sqrt{2}\,R$*).*

Proof. The proofs of Theorems 2.1. and 2.2. remain unchanged since they do not depend of the particular structure of the convex body G. In the proof of Theorem 2.3., the only change concerns the estimates in

Step 2 after Lemma 2.5. Instead of (2.35) – (2.44), it holds that

$$\Big|\frac{\partial y_i}{\partial s_j}(s)\Big| \leqslant \Big|\frac{\partial z_i}{\partial s_j}(s)\Big| + \Big|\frac{\partial y_i}{\partial s_j}(s) - \frac{\partial z_i}{\partial s_j}(s)\Big| \leqslant \Big|\frac{\partial z_i}{\partial s_j}(s)\Big| + C_1\,C_6\,\sigma_N \leqslant R + C_1\,C_6\,\sigma_N \quad \Longrightarrow \tag{2.63}$$

$$\operatorname*{Max}_{i,j}\Big|\frac{\partial y_i}{\partial s_j}(s)\Big| \leqslant R + C_1\,C_6\,\sigma_N \quad \Longrightarrow \quad \frac{R}{R + C_1\,C_6\,\sigma_N}\cdot \operatorname*{Max}_{i,j}\Big|\frac{\partial y_i}{\partial s_j}(s)\Big| \leqslant R\,. \tag{2.64}$$

Defining now $D \in (\,0\,,\,1\,)$ by

$$1 - D \,=\, \frac{R}{R + C_1\,C_6\,\sigma_N} \quad \Longrightarrow \quad D \,=\, \frac{C_1\,C_6}{R + C_1\,C_6\,\sigma_N}\cdot\sigma_N\,, \tag{2.65}$$

we may replace (2.43) and (2.44) by

$$\begin{aligned}
&\Big|\,z_i(s) - (1-D)\,y_i(s)\,\Big| \leqslant C_1\cdot C_5\cdot\sigma_N + D\cdot\Big|\,y_i(s)\,\Big| \\
&\leqslant C_1\,C_5 + \frac{C_1\,C_6}{R + C_1\,C_6\,\sigma_N}\cdot\sigma_N\cdot C_1\cdot\mathrm{Diam}\,(\Omega) \leqslant \Big(C_1\,C_5 + \frac{C_1\,C_6}{R}\cdot C_1\cdot\mathrm{Diam}\,(\Omega)\Big)\cdot\sigma_N \,=\, C_7\,\sigma_N\,; \qquad (2.66)\\
&\Big|\frac{\partial z_i}{\partial s_j}(s) - (1-D)\,\frac{\partial y_i}{\partial s_j}(s)\Big| \leqslant C_1\cdot C_6\cdot\sigma_N + D\Big|\frac{\partial y_i}{\partial s_j}(s)\Big| \\
&\qquad\qquad \leqslant \Big(C_1\,C_6 + \frac{C_1\,C_6}{R + C_1\,C_6\,\sigma_N}\cdot\sigma_N\cdot R \leqslant 2\,C_1\,C_6\,\sigma_N \,=\, C_8\,\sigma_n\,, \qquad (2.67)
\end{aligned}$$

and the proof may be completed as above. ∎

Remark. Note that the inequalities (2.33) and (2.34) have not been used in the proof of Corollary 2.7.

3. Application to the problem of edge detection within noisy image data.

a) The image denoising problem.

Since our approach for edge detection within noisy image data is closely connected with the classical image denoising procedure by variational methods, we start this section with a short introduction into the image denoising problem. In the following, we describe *greyscale images* through at least measurable functions x resp. I defined on a rectangle $\Omega \subset \mathbb{R}^2$ with values in $[\,0\,,\,1\,]$. The commonly used model for the capture of an original scene is the equation

$$I(s) = \mathcal{S}(\,x(s)\,) + \mathcal{N}(s) \tag{3.1}$$

where $x\colon\ \Omega \to [\,0\,,\,1\,]$ is an "ideal" image of the scene, the systematical errors (e. g. blur) are described by an operator $\mathcal{S}$, and $\mathcal{N}(s)$ is a noise term. [25] The formal solution of (3.1),

$$x(s) = \mathcal{S}^{-1}(\,I(s)\,) - \mathcal{S}^{-1}(\mathcal{N}(s)\,)\,, \tag{3.2}$$

however, leads to an ill-posed problem. Since the 90s, in generalization of the "scale-space" approach where the noisy image data $I(s)$ are subjected to a smoothing diffusion process, [26] the image denoising problem has been extensively treated by variational methods. The objectives within these problems minimize the defect $\mathcal{S}\big(\,x(s)\,\big) - I(s)$ together with a regularization term involving the first generalized partial derivatives

[25] Cf. [Aubert/Kornprobst 06], pp. 68 ff., and [Chambolle 00], pp. 7 ff.

[26] [Aubert/Kornprobst 06], pp. 94 ff., [Perona/Malik 90], [Weickert 96], pp. 2 – 18.

of x. For the diffusion methods, one must assume $x \in C^2(\Omega, \mathbb{R})$ whereas the variational problems can be formulated within Sobolev spaces $W^{1,p}(\Omega, \mathbb{R})$ resp. the space $BV(\Omega, \mathbb{R})$ [27] of the functions of bounded variation. [28] The original scenes x, however, in general even do not belong to $BV(\Omega, \mathbb{R})$, the most capacious of the mentioned spaces. [29] It must be emphasized that, for this reason, all methods proposed as yet are based on a compromise: instead of the original scene x, one searches for a representative with a prescribed grade of additional smoothness ("image smoothing"). In the following, we assume that $\mathcal{S}(\, x(s)\,) = x(s)$, and the variational problem will be formulated within the Sobolev space $W_0^{1,p}(\Omega, \mathbb{R})$:

$$(\mathrm{V})^{(1)}: \quad F(x) = \int_\Omega \big(\, x(s) - I(s)\,\big)^2\, ds \,+\, \mu \cdot \int_\Omega f(\,|\,\nabla x(s)\,|\,)\, ds \;\longrightarrow\; \mathrm{Min}\,!\,; \quad x \in W_0^{1,p}(\Omega, \mathbb{R}) \tag{3.3}$$

with $1 \leqslant p < \infty$, $I \in L^\infty(\Omega, \mathbb{R})$, $0 \leqslant I(s) \leqslant 1$ $(\forall)\, s \in \Omega$, a regularization parameter $\mu > 0$ and an integrand $f \in C^2(\mathbb{R}, \mathbb{R})$, which obeys an appropriate growth condition.
An important goal in image denoising/image smoothing is to conserve as far as possible the edge structure within the image data. [30] For this purpose, the use of an anisotropic regularization term is advisable. [31] In the following, we choose

$$\int_\Omega \sqrt{|\,\nabla x(s)\,|^2 + \eta^2}\; ds\,. \tag{3.4}$$

For sufficiently small values $\eta > 0$, this term may be understood as an approximation for the L^1- resp. total variation norm of ∇x but avoids its main disadvantages since the integrand in (3.4) is differentiable in $\mathfrak{o}$ and produces a reduced staircasing effect.

b) Edge detection within noisy image data by variational and optimal control methods.

Since Mumford/Shah's introduction of a functional including a measure for the total "length" of the edges within the given image data, [32] many efforts have been spent in the development of variational methods performing simultaneously both tasks of image restoration and edge detection. The commonly pursued strategy is to replace the cost functional in $(\mathrm{V})^{(1)}$ by a reasonable approximation for the Mumford-Shah functional, namely by a functional of so-called Ambrosio-Tortorelli type. [33] Besides the denoised image x, this functional depends on an additional variable e as a "sketch" for the edges with $e(s) \approx 0$ resp. $e(s) \approx 1$, depending on whether the point $s \in \Omega$ belongs to an edge within x or not. We obtain the following rather complicated variational problem:

$$\begin{aligned}(\mathrm{V})^{(2)}: \quad F(x,k) \;=\; & c_1(\varepsilon) \cdot \int_\Omega \big(\, x(s) - I(s)\,\big)^2\, ds \,+\, c_2(\varepsilon) \cdot \int_\Omega \big|\, \nabla x(s)\,\big|^p \cdot \big(\, e(s)^2 + c_4(\varepsilon)\,\big)\, ds \\ & + c_3(\varepsilon) \cdot \int_\Omega \Big(\, \varepsilon \cdot \big|\, \nabla e(s)\,\big|^2 + \frac{1}{4\,\varepsilon} \cdot \big(\, e(s) - 1\,\big)^2 \Big)\, ds \longrightarrow \mathrm{Min}\,!\,; \quad (x,e) \in W_0^{1,p}(\Omega, \mathbb{R}) \,\times\, W_0^{1,2}(\Omega, \mathbb{R})\end{aligned} \tag{3.5}$$

---

[27] Cf. [Evans/Gariepy 92], pp. 166 ff.

[28] For the treatment of image denoising problems within the space $BV(\Omega, \mathbb{R})$, we refer e. g. to [Aujol/Aubert/Blanc-Féraud/Chambolle 05], [Chambolle 00], [Chambolle/Lions 97], [Hinterberger/Hintermüller/Kunisch/Oehsen/Scherzer 03], [Osher/Burger/Goldfarb/Xu/Yin 05] and the seminal paper [Rudin/Osher/Fatemi 92].

[29] This conclusion will be suggested by [Gousseau/Morel 01].

[30] Here again, the inherent compromise within Sobolev space methods comes forward: "The theory seems to adopt again what it tried to avoid." ([Catté/Lions/Morel/Coll 92], p. 183).

[31] See [Aubert/Kornprobst 06], pp. 70 – 72.

[32] [Mumford/Shah 89], cf. also [Aubert/Kornprobst 06], pp. 153 ff.

[33] The approximation properties of the functional (3.5) have been studied in [Ambrosio/Tortorelli 92], p. 111, and [Bellettini/Coscia 94], p. 205, (2.1). See also [Aubert/Kornprobst 06], pp. 166 – 173.

with image data I as above, $\varepsilon > 0$ and weights $c_i(\varepsilon) > 0$, $1 \leqslant i \leqslant 4$. Within the objective, the first term is the classical fidelity term. The second one replaces the regularization term in $(\mathrm{V})^{(1)}$ and realizes a coupling between x and e, which favors values $e(s) \approx 0$ in points $s \in \Omega$ with large magnitudes of $\nabla x(s)$. Within the third term, the first member effects a quadratical regularization of e while the second member enforces $e(s) \approx 1$ except a subset of Ω of small measure. The interpretation of e as an edge detector is heuristically clear and may be rigorously justified by proving the Γ-convergence of the solutions of $(\mathrm{V})^{(2)}$ towards a solution of a variational problem with an Mumford-Shah type objective. [34)]
As an alternative to the study of $(\mathrm{V})^{(2)}$, however, we may add convex restrictions for ∇x to the problem $(\mathrm{V})^{(1)}$, thus converting the original variational problem into a multidimensional control problem of Dieudonné-Rashevsky type. Under comprehension of the regularization term (3.4), this problem reads as follows:

$$(\mathrm{P})^{(1)}: \quad F(x,u) = \int_\Omega \big(x(s) - I(s)\big)^2\, ds \,+\, \mu \cdot \int_\Omega \sqrt{u_1(s)^2 + u_2(s)^2 + \eta^2}\; ds \;\longrightarrow\; \mathrm{Min}\,! \,; \tag{3.6}$$

$$(x,u) \in W_0^{1,p}(\Omega,\mathbb{R}) \,\times\, L^p(\Omega,\mathbb{R}^2)\,; \tag{3.7}$$

$$\nabla x(s) \,=\, \begin{pmatrix} u_1(s) \\ u_2(s) \end{pmatrix} \quad (\forall)\, s \in \Omega\,; \tag{3.8}$$

$$u \in \mathrm{U} \,=\, \Big\{\, u \in L^p(\Omega,\mathbb{R}^2) \;\Big|\; |\,u_1(s)\,|^q \,+\, |\,u_2(s)\,|^q \,\leqslant\, R^q \;\; (\forall)\, s \in \Omega \,\Big\}\,. \tag{3.9}$$

Here we assume $1 \leqslant p < \infty$, $1 \leqslant q < \infty$, $\mu > 0$, $\eta > 0$, $R > 0$ and I as above. Note that, in consequence of (3.9), the feasible domain of $(\mathrm{P})^{(1)}$ lies in fact in $W_0^{1,\infty}(\Omega,\mathbb{R}) \times L^\infty(\Omega,\mathbb{R}^2)$. The edge detector e is immediately constructed from the control variables u_1 and u_2, e. g. [35)]

$$e(s) \,=\, 1 - \frac{1}{R^q}\Big(\,\big|\,u_1(s)\,\big|^q \,+\, \big|\,u_2(s)\,\big|^q\Big)\,. \tag{3.10}$$

Together with $(\mathrm{P})^{(1)}$, we consider the control problem $(\mathrm{P})^{(2)}$, which is identical but (3.9) is replaced by

$$u \in \mathrm{U} \,=\, \Big\{\, u \in L^p(\Omega,\mathbb{R}^2) \;\Big|\; \mathrm{Max}\big(\,|\,u_1(s)\,|\,,\,|\,u_2(s)\,|\,\big) \,\leqslant\, R \;\; (\forall)\, s \in \Omega \,\Big\}\,, \tag{3.11}$$

and the edge detector e is given by

$$e(s) \,=\, 1 - \frac{1}{R}\,\mathrm{Max}\big(\,|\,u_1(s)\,|\,,\,|\,u_2(s)\,|\,\big)\,. \tag{3.12}$$

In these optimal control reformulations of the problem, we interpret as "edges" those subsets of Ω where the control restriction becomes nearly active.
For $(\mathrm{P})^{(1)}$ and $(\mathrm{P})^{(2)}$, the assumptions of Theorem 2.1. are satisfied: Since $0 \leqslant I(s) \leqslant 1$ for all $s \in \Omega$ it follows that

$$\begin{aligned} &f(s,\xi,v) \,=\, \big(\xi - I(s)\big)^2 \,+\, \mu\,\sqrt{v_1^2 + v_2^2 + \eta^2} \\ &\Longrightarrow \quad \big|\,f(s,\xi,v)\,\big| \,\leqslant\, \xi^2 + 2\,|\,\xi\,| \,+\, \mu\,\sqrt{|\,v\,|^2 + \eta^2}\;, \end{aligned} \tag{3.13}$$

and the growth condition holds with $\varphi_1(s) \equiv 0$ and $\varphi_2(\,|\,\xi\,|\,,\,|\,v\,|\,) = |\,\xi\,|^2 + 2\,|\,\xi\,| + \mu\,\sqrt{|\,v\,|^2 + \eta^2}$.

[34)] [Bellettini/Coscia 94], p. 205 f., Theorem 2.1., for $p = 2$.

[35)] Cf. [Franek 07a], p. 65.

c) Numerical solution of the discretized control problem.

In the present paper, we pursue the second approach and solve the image denoising problem with simultaneous edge detection as a multidimensional control problem, applying the discretization method from Section 2 to $(\mathrm{P})^{(1)}$ and $(\mathrm{P})^{(2)}$. We choose $a = b = 2^N$ [36] and decompose $\Omega = [\,0\,,\,2^N\,]^2$ into $(K \times L)$ pixels $\mathrm{Q}_{k,l}$ with edge length 1 and northeastern vertex $s_{k,l}$ (consequently, $K = L = 2^N$). Then the mesh size amounts to $\sigma_N = \sqrt{2}/2$, and the noisy image data $I(s)$ are given as a pixelwise constant function. For this reason, the assumptions $(2.19) - (2.20)$ of Theorem 2.3., 3) are satisfied with $C_3 = 0$, and we may assume that $I(s) \,\big|\, \mathrm{Q}_{k,l} \equiv I(s_{k,l})$, $1 \leqslant k \leqslant K$, $1 \leqslant l \leqslant L$. Consequently, we state the following discretized problem:

$$(\mathrm{D})_N^{(1)}: \quad \widetilde{F}\big(\,\xi_{11}^{(1)},\,\ldots\,,\,\xi_{KL}^{(1)},\,v_{11}^{(1,1)},\,\ldots\,,\,v_{KL}^{(1,4)}\,\big) \;=\; \frac{1}{2}\cdot\frac{a\,b}{4^N}\cdot\sum_{k=1}^{K}\sum_{l=1}^{L}\Big(\,\big(\,\xi_{k-1,l-1}^{(1)} - I(s_{k,l})\,\big)^2 \;+\; \big(\,\xi_{k,l}^{(1)} - I(s_{k,l})\,\big)^2$$
$$+\,\mu\,\sqrt{\big(\,v_{k,l}^{(1,1)}\,\big)^2 + \big(\,v_{k,l}^{(1,2)}\,\big)^2 + \eta^2} \;+\; \mu\,\sqrt{\big(\,v_{k,l}^{(1,3)}\,\big)^2 + \big(\,v_{k,l}^{(1,4)}\,\big)^2 + \eta^2}\,\Big) \;\longrightarrow\; \mathrm{Min}\,! \,; \tag{3.14}$$

$$\big(\,\xi_{0,0}^{(1)},\,\ldots,\,\xi_{K,L}^{(1)},\,v_{1,1}^{(1,1)},\,\ldots,\,v_{K,L}^{(1,4)}\,\big) \;\in\; \mathbb{R}^{(K+1)(L+1)} \times \mathbb{R}^{4KL}\,; \tag{3.15}$$

$$\xi_{0,l}^{(1)} \;=\; \xi_{K,l}^{(1)} \;=\; 0\,; \qquad\qquad \xi_{k,0}^{(1)} \;=\; \xi_{k,L}^{(1)} \;=\; 0\,,\;\; 1 \leqslant k \leqslant K\,,\;\; 1 \leqslant l \leqslant L\,; \tag{3.16}$$

$$v_{k,l}^{(1,1)} \;=\; \frac{\xi_{k,l-1}^{(1)} - \xi_{k-1,l-1}^{(1)}}{(a/2^N)}\,; \quad v_{k,l}^{(1,2)} \;=\; \frac{\xi_{k,l}^{(1)} - \xi_{k,l-1}^{(1)}}{(b/2^N)}\,,\;\; 1 \leqslant k \leqslant K\,,\;\; 1 \leqslant l \leqslant L\,; \tag{3.17}$$

$$v_{k,l}^{(1,3)} \;=\; \frac{\xi_{k,l}^{(1)} - \xi_{k-1,l}^{(1)}}{(a/2^N)}\,; \qquad v_{k,l}^{(1,4)} \;=\; \frac{\xi_{k-1,l}^{(1)} - \xi_{k-1,l-1}^{(1)}}{(b/2^N)}\,,\;\; 1 \leqslant k \leqslant K\,,\;\; 1 \leqslant l \leqslant L\,; \tag{3.18}$$

$$\big|\,v_{k,l}^{(1,1)}\,\big|^q \;+\; \big|\,v_{k,l}^{(1,2)}\,\big|^q \;\leqslant\; R^q\,,\;\; 1 \leqslant k \leqslant K\,,\;\; 1 \leqslant l \leqslant L\,; \tag{3.19}$$

$$\big|\,v_{k,l}^{(1,3)}\,\big|^q \;+\; \big|\,v_{k,l}^{(1,4)}\,\big|^q \;\leqslant\; R^q\,,\;\; 1 \leqslant k \leqslant K\,,\;\; 1 \leqslant l \leqslant L\,. \tag{3.20}$$

As the discretization of the edge detector, we obtain

$$e(s_{k,l}) \;=\; 1 - \frac{1}{R^q}\cdot\mathrm{Max}\,\Big(\,\big|\,v_{k,l}^{(1,1)}\,\big|^q \;+\; \big|\,v_{k,l}^{(1,2)}\,\big|^q\,,\,\big|\,v_{k,l}^{(1,3)}\,\big|^q \;+\; \big|\,v_{k,l}^{(1,4)}\,\big|^q\Big)\,. \tag{3.21}$$

Accordingly, in $(\mathrm{D})_N^{(2)}$ as the dicretization of $(\mathrm{P})^{(2)}$, (3.19) and (3.20) will be replaced by

$$\big|\,v_{k,l}^{(1,1)}\,\big| \;\leqslant\; R\,,\;\; \big|\,v_{k,l}^{(1,2)}\,\big| \;\leqslant\; R\,,\;\; \big|\,v_{k,l}^{(1,3)}\,\big| \;\leqslant\; R\,,\;\; \big|\,v_{k,l}^{(1,4)}\,\big| \;\leqslant\; R\,,\;\; 1 \leqslant k \leqslant K\,,\;\; 1 \leqslant l \leqslant L\,, \tag{3.22}$$

and the edge detector (3.12) will be discretized as

$$e(s_{k,l}) \;=\; 1 - \frac{1}{R}\cdot\mathrm{Max}\,\Big(\,\big|\,v_{k,l}^{(1,1)}\,\big|\,,\,\big|\,v_{k,l}^{(1,2)}\,\big|\,,\,\big|\,v_{k,l}^{(1,3)}\,\big|\,,\,\big|\,v_{k,l}^{(1,4)}\,\big|\,\Big)\,. \tag{3.23}$$

The evaluation of the necessary optimality conditions (Karush-Kuhn-Tucker conditions) for $(\mathrm{D})_N^{(1)}$ results in a large-scale system of nonlinear equations and inequalities, which may be solved with high precision and efficiency by interior point methods. As input/output platform, MATLAB has been used while the discretized problem has been formulated with the aid of the modelling language AMPL [37] and then transferred to the

[36] We use $N = 7$ in the experiments with the Lena, Peppers and Cameraman images and $N = 9$ in the experiments with the Lighthouse image.

[37] AMPL is a commercially distributed modelling language, which allows for the description of large-scale optimization problems, the interaction with a solver and a further processing of the output data. Cf. [Fourer/Gay/Kernighan 03] .

interior-point solver IPOPT. [38] The results have been represented, evaluated and archived with MATLAB again.

d) Image data and evaluation of the results.

For our numerical experiments, we used segments I of some well-known test images (Lena, [39] Peppers, [40] Cameraman [41]) with $K = L = 128$ with or without artificial addition of white noise [42] (Figs. 3.1. and 3.17., 3.21. and 3.25., 3.29. and 3.33.) as well as the Lighthouse image [43] with $K = L = 512$ (Fig. 3.37.). The quality of the image restoration x will be evaluated by the SNR indicator

$$SNR\,(x, I) \;=\; -10 \log_{10} \Big(\frac{\displaystyle\int_\Omega \big(x(s) - I(s) \big)^2 \, ds}{\displaystyle\int_\Omega x(s)^2 \, ds} \Big) \;\approx\; -10 \log_{10} \Big(\frac{\displaystyle\sum_{k=4}^{K-3} \sum_{l=4}^{L-3} \big(x_{kl} - I(s_{kl}) \big)^2}{\displaystyle\sum_{k=4}^{K-3} \sum_{l=4}^{L-3} \big(x_{kl} \big)^2} \Big) \tag{3.24}$$

together with the relative error of the reconstruction x with respect to the original image $\hat{x}$ [44]

$$E_{rel}\,(x, \hat{x}) \;=\; 100 \cdot \left(\frac{\displaystyle\int_\Omega \big(x(s) - \hat{x}(s) \big)^2 \, ds}{\displaystyle\int_\Omega \hat{x}(s)^2 \, ds} \right)^{1/2} \;\approx\; 100 \cdot \left(\frac{\displaystyle\sum_{k=4}^{K-3} \sum_{l=4}^{L-3} \big(x_{kl} - \hat{x}(s_{kl}) \big)^2}{\displaystyle\sum_{k=4}^{K-3} \sum_{l=4}^{L-3} \big(\hat{x}(s_{kl}) \big)^2} \right)^{1/2} . \tag{3.25}$$

For the quantitative evaluation of the edge sketches e, however, the literature does not provide a generally accepted criterion as yet. For this reason, we propose the following error measure IEE ("intensity edge error"): [45]

$$IEE\,(e, \hat{e}) \;=\; 100 \cdot \left(\frac{\displaystyle\int_\Omega \big(e(s) - \hat{e}(s) \big)^2 \, ds}{\displaystyle\int_\Omega \hat{e}(s)^2 \, ds} \right)^{1/2} \;\approx\; 100 \cdot \left(\frac{\displaystyle\sum_{k=4}^{K-3} \sum_{l=4}^{L-3} \big(e(s_{kl}) - \hat{e}(s_{kl}) \big)^2}{\displaystyle\sum_{k=4}^{K-3} \sum_{l=4}^{L-3} \big(\hat{e}(s_{kl}) \big)^2} \right)^{1/2} \tag{3.26}$$

with reference to the "ideal" edge sketches $\hat{e}$ (Figs. 3.2., 3.22., 3.30. and 3.38.), which have been obtained by application of EDISON, a non-variational standard method for edge detection, to the (noiseless) original image data. The parameters within EDISON have been chosen in accordance with [Meer/Georgescu 01], p. 1361, Fig. 10 (c), where the application of the algorithm to the Cameraman image has been documented. [46] In all cases, a margin of 3 pixels width has been excluded from the evaluation.

---

[38] Cf. [Laird/Wächter 09] and [Wächter/Biegler 06]. The experiments have been performed with version 3.6.1., compiled with the MA27 routine.

[39] Accessible via http://www.dcs.qmul.ac.uk/∼phao/CIP/Images/Lena.tif (last access: 16. 02. 2009).

[40] Accessible via http://www.dcs.qmul.ac.uk/∼phao/CIP/Images/Peppers.tif (last access: 16. 02. 2009).

[41] Accessible via http://vip.cs.utsa.edu/classes/cs6723f2001/lectures/images/cameraman.tif (last access: 16. 02. 2009).

[42] In all cases with standard deviation zero and variance $\sigma = 0.01$.

[43] Accessible via http://www.hlevkin.com/TestImages/lighthouse.bmp (last access: 16. 02. 2009).

[44] This error measure may be interpreted as well in terms of a "method noise", cf. [Buades/Coll/Morel 05], pp. 492 ff.

[45] Cf. [Brune/Maurer/Wagner 09], p. 1199, Definition 4.2.

[46] The following parameters have been used: Fig. 3.2.: grad win. 3; min. length 10; nonmaxima suppr.: arc / 0.7 / 0.5; hysteresis thresh. low: arc / 0.7 / 0.8; hysteresis thresh. high: box / 0.93 / 0.93. Fig. 3.22.: grad win. 3; min. length 10; nonmaxima suppr.: arc / 0.7 / 0.5; hysteresis thresh. low: arc / 0.7 / 0.8; hysteresis thresh. high: box / 0.9 / 0.9. Figs. 3.30. and 3.38.: grad win. 2; min. length 10; nonmaxima suppr.: arc / 0.7 / 0.5; hysteresis thresh. low: arc / 0.7 / 0.8; hysteresis thresh. high: box / 0.9 / 0.9.

e) Numerical results.

In Figs. 3.3. – 3.20., we present our results for the Lena image. In a first series (Figs. 3.3. – 3.16.), we show experiments with the original Lena image from Fig. 3.1. In order to compare our optimal control approach with the Ambrosio-Tortorelli method, the variational problem $(\mathrm{V})^{(2)}$ has been solved first (Figs. 3.3. and 3.4.). [47] In Figs. 3.5. – 3.12., the solutions for the multidimensional control problem $(\mathrm{P})^{(1)}$ are depicted. Here we take $p = q = 2$, $\mu = 0.05$ and vary the parameter R. Then we solve $(\mathrm{P})^{(2)}$ with $p = 2$ and $\mu = 0.1$ and show again results for different values of R (Figs. 3.13. – 3.16.). The second series has been calculated with the noisy Lena image from Fig. 3.17. We present three edge sketches: the first one generated by EDISON (using the same parameters as in Fig. 3.2.), the second one calculated with the Ambrosio-Tortorelli method, and the third one a typical result obtained by the optimal control method.
For the Peppers and the Cameraman image, the edge sketches for the original and noisy images generated by EDISON (Figs. 3.22. and 3.26, 3.30. and 3.34.) and the Ambrosio-Tortorelli method (Figs. 3.23. and 3.27., 3.31. and 3.35.) will be shown together with one typical result of the optimal control method in each case (Figs. 3.24. and 3.28., 3.32. and 3.36.) ($p = q = 2$).
The Lighthouse image has been selected in order to show exemplarily how the transition from structure to texture will be reflected in the output of the optimal control method. For this purpose, we present two edge sketches for original image (Figs. 3.39. and 3.40.) calculated with the optimal control method based on the problems $(\mathrm{P})^{(1)}$ and $(\mathrm{P})^{(2)}$ with $p = q = 2$, respectively.

Fig. 3.1.

Fig. 3.2.

Left: Segment $\hat{x}$ of the original Lena image.
Right: Reference edge sketch $\hat{e}$ for the original Lena image, generated with EDISON.

Fig. 3.3.

Fig. 3.4.

Variational problem $(\mathrm{V})^{(2)}$ for the original Lena image.
Left: Restored image x.
Right: Edge sketch e.
$p = 2$, $\varepsilon = 0.125$, $c_1(\varepsilon) = 800$, $c_2(\varepsilon) = 30$, $c_3(\varepsilon) = 0.2$, $c_4(\varepsilon) = 0$,
$IEE\,(e, \hat{e}) = 28.9651$,
$E_{rel}\,(x, \hat{x}) = 0.1679$,
$SNR\,(x, \hat{x}) = 55.4932$.

[47] Analogously to [Bourdin 99], a direct method has been applied to $(\mathrm{V})^{(2)}$ as well.

Fig. 3.5.

Fig. 3.6.

Control problem $(\mathrm{P})^{(1)}$ for the original Lena image.
Left: Restored image x.
Right: Edge sketch e.
$\mu = 0.05$, $\eta = 0.003$, $R = 0.35$,
$IEE\,(e, \hat{e}) = 29.4748$,
$E_{rel}\,(x, \hat{x}) = 5.3760$,
$SNR\,(x, \hat{x}) = 25.3714$.

Fig. 3.7.

Fig. 3.8.

Control problem $(\mathrm{P})^{(1)}$ for the original Lena image.
Left: Restored image x.
Right: Edge sketch e.
$\mu = 0.05$, $\eta = 0.003$, $R = 0.25$,
$IEE\,(e, \hat{e}) = 28.8554$,
$E_{rel}\,(x, \hat{x}) = 5.4930$,
$SNR\,(x, \hat{x}) = 25.1812$.

Fig. 3.9.

Fig. 3.10.

Control problem $(\mathrm{P})^{(1)}$ for the original Lena image.
Left: Restored image x.
Right: Edge sketch e.
$\mu = 0.05$, $\eta = 0.003$, $R = 0.20$,
$IEE\,(e, \hat{e}) = 29.5749$,
$E_{rel}\,(x, \hat{x}) = 5.6808$,
$SNR\,(x, \hat{x}) = 24.8835$.

Fig. 3.10.

Fig. 3.12.

Control problem $(\mathrm{P})^{(1)}$ for the original Lena image.
Left: Restored image x.
Right: Edge sketch e.
$\mu = 0.05$, $\eta = 0.005$, $R = 0.15$,
$IEE\,(e, \hat{e}) = 33.2925$,
$E_{rel}\,(x, \hat{x}) = 6.6966$,
$SNR\,(x, \hat{x}) = 23.4163$.

Fig. 3.13.

Fig. 3.14.

Control problem $(\mathrm{P})^{(2)}$ for the original Lena image.
Left: Restored image x.
Right: Edge sketch e.
$\mu = 0.1$, $\eta = 0.005$, $R = 0.275$,
$IEE\,(e, \hat{e}) = 26.8231$,
$E_{rel}\,(x, \hat{x}) = 6.2591$,
$SNR\,(x, \hat{x}) = 24.0272$.

Fig. 3.15.

Fig. 3.16.

Control problem $(\mathrm{P})^{(2)}$ for the original Lena image.
Left: Restored image x.
Right: Edge sketch e.
$\mu = 0.1$, $\eta = 0.005$, $R = 0.15$,
$IEE\,(e, \hat{e}) = 30.9092$,
$E_{rel}\,(x, \hat{x}) = 7.1298$,
$SNR\,(x, \hat{x}) = 22.8532$.

Fig. 3.17.

Fig. 3.18

Left: Segment $\hat{x}$ of the noisy Lena image.
Right: Edge sketch e for the noisy Lena image, generated with EDISON (parameters as in Fig. 3.2.).

Fig. 3.19.

Fig. 3.20.

Left: Variational problem $(\mathrm{V})^{(2)}$: edge sketch e for the noisy Lena image (restored image x not shown).
$p = 2$, $\varepsilon = 0.5$, $c_1(\varepsilon) = 5$, $c_2(\varepsilon) = 10$,
$c_3(\varepsilon) = 0.3$, $c_4(\varepsilon) = 0.01$
$IEE\,(e, \hat{e}) = 30.0624$,
$E_{rel}\,(x, \hat{x}) = 9.7103$, $SNR\,(x, \hat{x}) = 18.5300$.
Right: Control problem $(\mathrm{P})^{(2)}$ for the noisy Lena image: edge sketch e (restored image x not shown).
$\mu = 0.40$, $\eta = 0.005$, $R = 0.25$,
$IEE\,(e, \hat{e}) = 28.4887$,
$E_{rel}\,(x, \hat{x}) = 9.9958$, $SNR\,(x, \hat{x}) = 14.0717$.

Fig. 3.21.

Fig. 3.22.

Left: Segment $\hat{x}$ of the original Peppers image.
Right: Reference edge sketch $\hat{e}$ for the original Peppers image, generated with EDISON.

Fig. 3.23.

Fig. 3.24.

Left: Variational problem $(\mathrm{V})^{(2)}$ for the original Peppers image: edge sketch e (restored image x not shown).
$p = 2$, $\varepsilon = 0.125$, $c_1(\varepsilon) = 800$, $c_2(\varepsilon) = 30$, $c_3(\varepsilon) = 0.5$, $c_4(\varepsilon) = 0$,
$IEE\,(e, \hat{e}) = 28.4906$,
$E_{rel}\,(x, \hat{x}) = 0.2310$, $SNR\,(x, \hat{x}) = 52.7222$.
Right: Control problem $(\mathrm{P})^{(1)}$ for the original Peppers image: edge sketch e (restored image x not shown).
$\mu = 0.05$, $\eta = 0.003$, $R = 0.25$,
$IEE\,(e, \hat{e}) = 28.6631$,
$E_{rel}\,(x, \hat{x}) = 8.0746$, $SNR\,(x, \hat{x}) = 21.7975$.

Fig. 3.25.

Fig. 3.26.

Left: Noisy Peppers image.
Right: Edge sketch e for the noisy Peppers image, generated with EDISON (parameters as in Fig. 3.22.).

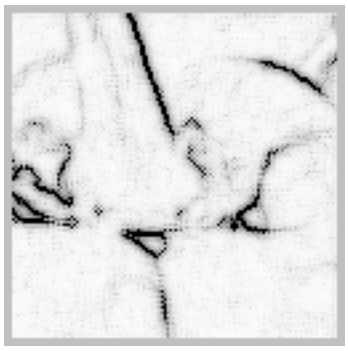

Fig. 3.27.

Fig. 3.28.

Left: Variational problem $(\mathrm{V})^{(2)}$: edge sketch e for the noisy Peppers image (restored image x not shown).
$p = 2$, $\varepsilon = 0.5$, $c_1(\varepsilon) = 3$, $c_2(\varepsilon) = 10$, $c_3(\varepsilon) = 0.2$, $c_4(\varepsilon) = 0$,
$IEE\,(e, \hat{e}) = 30.5241$,
$E_{rel}\,(x, \hat{x}) = 9.8737$, $SNR\,(x, \hat{x}) = 16.9203$.
Right: Control problem $(\mathrm{P})^{(1)}$ for the noisy Peppers image: edge sketch e (restored image x not shown).
$\mu = 0.15$, $\eta = 0.003$, $R = 0.25$,
$IEE\,(e, \hat{e}) = 28.9618$,
$E_{rel}\,(x, \hat{x}) = 9.8851$, $SNR\,(x, \hat{x}) = 14.3100$.

Fig. 3.29.

Fig. 3.30.

Left: Segment $\hat{x}$ of the original Cameraman image.
Right: Reference edge sketch $\hat{e}$ for the original Cameraman image, generated with EDISON.

Fig. 3.31.

Fig. 3.32.

Left: Variational problem $(\mathrm{V})^{(2)}$ for the original Cameraman image: edge sketch e (restored image x not shown).
$p = 2$, $\varepsilon = 0.125$, $c_1(\varepsilon) = 800$, $c_2(\varepsilon) = 10$, $c_3(\varepsilon) = 0.4$, $c_4(\varepsilon) = 0.01$,
$IEE\,(e, \hat{e}) = 27.4784$,
$E_{rel}\,(x, \hat{x}) = 0.1317$, $SNR\,(x, \hat{x}) = 57.6020$.
Right: Control problem $(\mathrm{P})^{(2)}$ for the original Cameraman image: edge sketch e (restored image x not shown).
$\mu = 0.17$, $\eta = 0.015$, $R = 0.425$,
$IEE\,(e, \hat{e}) = 27.6191$,
$E_{rel}\,(x, \hat{x}) = 16.3432$, $SNR\,(x, \hat{x}) = 15.4916$.

Fig. 3.33.

Fig. 3.34.

Left: Noisy Cameraman image.
Right: Edge sketch for the noisy Cameraman image, generated with EDISON (parameters as in Fig. 3.30.).

Fig. 3.35.

Fig. 3.36.

Left: Variational problem $(\mathrm{V})^{(2)}$: edge sketch e for the noisy Cameraman image (restored image x not shown).
$p = 2$, $\varepsilon = 0.25$, $c_1(\varepsilon) = 5$, $c_2(\varepsilon) = 10$, $c_3(\varepsilon) = 0.5$, $c_4(\varepsilon) = 0$,
$IEE\,(e, \hat{e}) = 29.9350$,
$E_{rel}\,(x, \hat{x}) = 16.3001$, $SNR\,(x, \hat{x}) = 15.1594$.
Right: Control problem $(\mathrm{P})^{(2)}$ for the noisy Cameraman image: edge sketch e (restored image x not shown).
$\mu = 0.17$, $\eta = 0.012$, $R = 0.44$,
$IEE\,(e, \hat{e}) = 27.8413$,
$E_{rel}\,(x, \hat{x}) = 17.4697$, $SNR\,(x, \hat{x}) = 12.0828$.

Fig. 3.37.

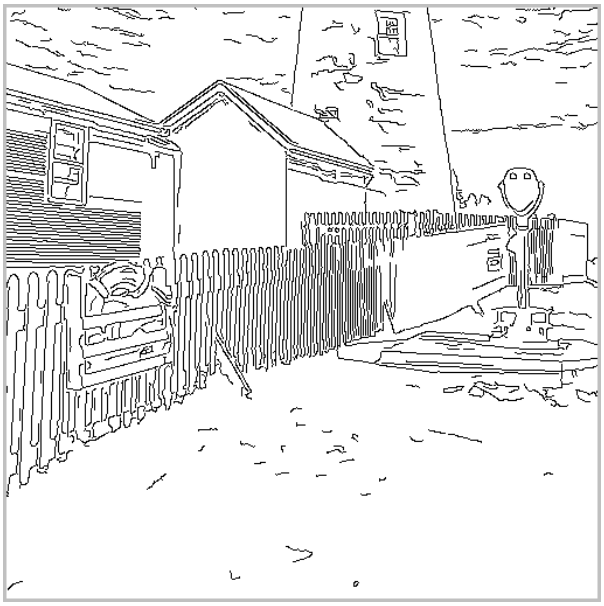

Fig. 3.38.

Left: Original Lighthouse image $\hat{x}$.
Right: Reference edge sketch $\hat{e}$ for the original Lighthouse image, generated with EDISON.

Fig. 3.39.

Fig. 3.40.

Left: Control problem $(\mathrm{P})^{(1)}$ for the original Lighthouse image: edge sketch e generated with
$\mu = 0.05$, $\eta = 0.005$, $R = 0.30$,
$IEE\,(e, \hat{e}) = 25.9636$, $E_{rel}\,(x, \hat{x}) = 11.5812$, $SNR\,(x, \hat{x}) = 18.6371$.
Right: Control problem $(\mathrm{P})^{(2)}$ for the original Lighthouse image: edge sketch e generated with
$\mu = 0.11$, $\eta = 0.01$, $R = 0.275$,
$IEE\,(e, \hat{e}) = 24.4521$, $E_{rel}\,(x, \hat{x}) = 12.4732$, $SNR\,(x, \hat{x}) = 17.9665$.

f) Discussion of the results.

A *visual comparison* of the results must take into account that a non-variational edge detection method like EDISON generates a binary sketch with edges of constant width while the output produced by the Ambrosio-Tortorelli as well as by the optimal control method is a greyscale image showing edges of different width and intensity. If one wishes to use the edge sketch as a representation of the geometrical information contained in the image, the second class of methods is clearly preferable.
The comparison between the variational and optimal control method shows that the edge sketches supplied by the optimal control method look somewhat clearer than those generated by the Ambrosio-Tortorelli method, which has a tendency to produce minute grainy items. Both methods are able to resolve fine details, e. g. the fringe at the margin of the hat in the Lena image and the white lever in the center of the Cameraman image.[48] Particularly in application to the noisy images, the optimal control method shows a moderate staircasing effect (which will be stronger in $(\mathrm{P})^{(2)}$ than in $(\mathrm{P})^{(1)}$, compare Fig. 3.12. with Fig. 3.16.). Further, the optimal control method tends to suppress small details and texture elements automatically, while non-variational methods like EDISON must perform some postprocessing steps in order to achieve the same goal. The experiments with the Lighthouse image (Figs. 3.39. and 3.40.) provide a closer insight into the different recognization of structures and textures (e. g. the lawn in the foreground, the timber wall and the fence) by the optimal control method. As one may expect from the choice of the regularization term, the restored versions of the images provided by the optimal control method look somewhat cartoon-like.
In order to perform a *quantitative comparison*, the best IEE values obtained by the respective methods when applied to noisy data have been collected in the following table.

	Lena (Fig. 3.17.)	Peppers (Fig. 3.25.)	Cameraman (Fig. 3.33.)
EDISON	30.0912 (Fig. 3.18.)	28.4475 (Fig. 3.26.)	26.6190 (Fig. 3.34.)
$(\mathrm{V})^{(2)}$	30.0624 (Fig. 3.19.)	30.5241 (Fig. 3.27)	29.9350 (Fig. 3.35.)
$(\mathrm{P})^{(1)}$	29.2798 (not shown)	28.9618 (Fig. 3.28.)	28.9875 (not shown)
$(\mathrm{P})^{(2)}$	28.4887 (Fig. 3.20.)	27.9649 (not shown)	27.8413 (Fig. 3.36.)

At a single glance, one realizes that the results of the three methods are in a comparable range. A closer inspection reveals a trend of increasing quality of edge detection in the order $(\mathrm{V})^{(2)} \to (\mathrm{P})^{(1)} \to (\mathrm{P})^{(2)}$, which will be supported by the experiments with the original images as well. It comes as a little surprise that, for two of the three test images, the optimal control method provides even better results than the EDISON edge generator. For the noisy images, the SNR and E_{rel} values produced by the variational and the optimal control method are comparable as well.
With respect to the *variation of* R within the optimal control method, the best results both with respect to the visual comparison as well as for the IEE criterion have been obtained for $R \approx 0.3$. In many respects, the sharpness R of the control restriction behaves like an additional regularization parameter. If the value of R is selected too small then "overcrowded" edge sketches develop (Figs. 3.12. und 3.16.). Although the assumptions of the convergence theorem for $(\mathrm{P})^{(1)}$ (Theorem 2.3.) are only satisfied for

$$R < \mathrm{Min}\left(\frac{1}{2}\,,\,\frac{\sqrt{3}-\sqrt{2}}{2\sqrt{2}}\right) = 0.112372\ldots\,, \tag{3.27}$$

reasonable results have been obtained even for $R \in [\,0.11\,,\,0.4\,]$. For the Lena image, an enlargement of R beyond 0.35 results in "bleached out" edge sketches quite similar to Fig. 3.6. (not depicted).

[48] Compare with [Chan/Esedoḡlu/Nikolova 06], p. 1645 f., Figs. 7 and 9, where this detail cannot be seen.

A drawback shared by all investigated methods is the *necessity for the empirical determination of the parameters* within the objectives. In the optimal control approach, however, we must deal essentially with the two parameters μ and R (since p, q and η may be reasonably fixed in advance), while the Ambrosio-Tortorelli method and EDISON involve the simultaneous adaptation of five or eight parameters, respectively.
To close with a few remarks concerning *compatibility, implementation and runtime behaviour*, let us note first that, in the non-variational method, the edge detection is always performed as a separate step, possibly followed by additional postprocessing steps. In contrast, the variational as well as the optimal control method allow to combine the edge detection with further optimization goals. In comparison with the Ambrosio-Tortorelli method, the optimal control method has the advantage that only half the number of variables must be implemented, replacing the control variables $v_{k,l}^{(1,1)}$, ... , $v_{k,l}^{(1,4)}$ by finite differences according to (3.17) and (3.18). For large image data (e. g. the Lighthouse image with $N = 512$), this fact results in a considerable speedup. Even for $N = 128$ and noisy image data, the optimal control method runs definitely faster than the variational method. Consequently, if edge detection is to be integrated as a subtask into a more comprehensive problem, e. g. multimodal image registration, the optimal control method should be preferred.

g) Conclusion.

We may summarize that the three presented edge detection methods supply results of comparable quality. Consequently, our experiments demonstrate that the treatment of the image denoising problem with simultaneous edge detection as a multidimensional control problem offers a real alternative to the existing variational methods and even to the application of non-variational edge detectors.

Acknowledgement.

Since September 2009, Marcus Wagner has been supported within the Special Research Unit "Mathematical Optimization and Applications in Biomedical Sciences" (Graz) by the Austrian Science Fund.

References.

1. [Adams/Fournier 07] Adams, R. A.; Fournier, J. J. F.: *Sobolev Spaces.* Academic Press / Elsevier; Amsterdam etc. 2007, 2. Aufl.
2. [Ambrosio/Tortorelli 92] Ambrosio, L.; Tortorelli, V. M.: *On the approximation of free discontinuity problems.* Boll. Un. Mat. Ital. B (7) **6** (1992), 105 – 123
3. [Andrejewa/Klötzler 84a] Andrejewa, J. A.; Klötzler, R.: *Zur analytischen Lösung geometrischer Optimierungsaufgaben mittels Dualität bei Steuerungsproblemen. Teil I.* Z. Angew. Math. Mech. **64** (1984), 35 – 44
4. [Andrejewa/Klötzler 84b] Andrejewa, J. A.; Klötzler, R.: *Zur analytischen Lösung geometrischer Optimierungsaufgaben mittels Dualität bei Steuerungsproblemen. Teil II.* Z. Angew. Math. Mech. **64** (1984), 147 – 153
5. [Aujol/Aubert/Blanc-Féraud/Chambolle 05] Aujol, J.-F.; Aubert, G.; Blanc-Féraud, L.; Chambolle, A.: *Image decomposition into a bounded variation component and an oscillating component.* J. Math. Imaging Vision **22** (2005), 71 – 88
6. [Aubert/Kornprobst 06] Aubert, G.; Kornprobst, P.: *Mathematical Problems in Image Processing: Partial Differential Equations and the Calculus of Variations.* Springer; New York etc. 2006, 2nd ed.
7. [Bellettini/Coscia 94] Bellettini, G.; Coscia, A.: *Discrete approximation of a free discontinuity problem.* Numer. Funct. Anal. Optim. **15** (1994), 201 – 224
8. [Bourdin 99] Bourdin, B.: *Image segmentation with a finite element method.* M2AN Mathematical Modelling and Numerical Analysis **33** (1999), 229 – 244

9. [Brokate 85] Brokate, M.: *Pontryagin's principle for control problems in age-dependent population dynamics.* J. Math. Biology **23** (1985), 75 – 101
10. [Brune/Maurer/Wagner 09] Brune, C.; Maurer, H.; Wagner, M.: *Detection of intensity and motion edges within optical flow via multidimensional control.* SIAM J. Imaging Sci. **2** (2009), 1190 – 1210
11. [Buades/Coll/Morel 05] Buades, A.; Coll, B.; Morel, J.-M.: *A review of image denoising algorithms, with a new one.* Multiscale Model Simul. **4** (2005), 490 – 530
12. [Catté/Lions/Morel/Coll 92] Catté, F.; Lions, P.-L.; Morel, J.-M.; Coll, T.: *Image selective smoothing and edge detection by nonlinear diffusion.* SIAM J. Numer. Anal. **29** (1992), 182 – 193
13. [Chambolle 00] Chambolle, A.: *Mathematical problems in image processing. Inverse problems in image processing and image segmentation: some mathematical and numerical aspects.* ICTP Lecture Notes, II. Abdus Salam International Centre for Theoretical Physics, Trieste 2000 (electronic)
14. [Chambolle/Lions 97] Chambolle, A.; Lions, P.-L.: *Image recovery via total variation minimization and related problems.* Numer. Math. **76** (1997), 167 – 188
15. [Chan/Esedoḡlu/Nikolova 06] Chan, T. F.; Esedoḡlu, S.; Nikolova, M.: *Algorithms for finding global minimizers of image segmentation and denoising models.* SIAM J. Appl. Math. **66** (2006), 1632 – 1648
16. [Chan/Vese 01] Chan, T. F.; Vese, L. : *Active contours without edges.* IEEE Trans. Image Processing **10** (2001), 266 – 277
17. [Ciarlet 87] Ciarlet, P.: *The Finite Element Methods for Elliptic Problems.* North-Holland; Amsterdam - New York - Oxford - Tokyo 1987, 2nd ed.
18. [Dacorogna 08] Dacorogna, B.: *Direct Methods in the Calculus of Variations.* Springer; New York etc. 2008, 2nd ed.
19. [Dacorogna/Marcellini 97] Dacorogna, B.; Marcellini, P.: *General existence theorems for Hamilton-Jacobi equations in the scalar and vectorial case.* Acta Mathematica **178** (1997), 1 – 37
20. [Dacorogna/Marcellini 98] Dacorogna, B.; Marcellini, P.: *Cauchy-Dirichlet problem for first order nonlinear systems.* J. Funct. Anal. **152** (1998), 404 – 446
21. [Dacorogna/Marcellini 99] Dacorogna, B.; Marcellini, P.: *Implicit Partial Differential Equations.* Birkhäuser; Boston - Basel - Berlin 1999
22. [Dewess/Helbig 95] Deweß, G.; Helbig, P.: *Einschätzen und Optimieren von Verkehrsflüssen.* In: Bachem, A.; Jünger, M.; Schrader, R. (Eds.): Mathematik in der Praxis. Fallstudien aus Industrie, Wirtschaft, Naturwissenschaften und Medizin. Springer; Berlin etc. 1995, 473 – 492
23. [Droske/Rumpf 04] Droske, M.; Rumpf, M.: *A variational approach to nonrigid morphological image registration.* SIAM J. Appl. Math. **64** (2004), 668 – 687
24. [Ekeland/Témam 99] Ekeland, I.; Témam, R.: *Convex Analysis and Variational Problems.* SIAM; Philadelphia 1999, 2nd ed.
25. [Evans/Gariepy 92] Evans, L. C.; Gariepy, R. F.: *Measure Theory and Fine Properties of Functions.* CRC Press; Boca Raton etc. 1992
26. [Faugeras/Hermosillo 04] Faugeras, O.; Hermosillo, G.: *Well-posedness of two nonrigid multimodal image registration methods.* SIAM J. Appl. Math. **64** (2004), 1550 – 1587
27. [Feichtinger/Tragler/Veliov 03] Feichtinger, G.; Tragler, G.; Veliov, V. M.: *Optimality conditions for age-structured control systems.* J. Math. Anal. Appl. **288** (2003), 47 – 68
28. [Felzenszwalb/Huttenlocher 04] Felzenszwalb, P. F.; Huttenlocher, D. P.: *Efficient graph-based image segmentation.* Int. J. Computer Vision **59** (2004), 167 – 181
29. [Fourer/Gay/Kernighan 03] Fourer, R.; Gay, D. M.; Kernighan, B. W.: *AMPL: A Modeling Language for Mathematical Programming.* Brooks/Cole — Thomson Learning; Pacific Grove 2003, 2nd ed.
30. [Franek 07a] Franek, L.: *Anwendung optimaler Steuerungsprobleme mit L^∞-Steuerbeschränkung auf ein Modellproblem der Bildverarbeitung.* Diploma thesis; University of Münster 2007
31. [Franek 07b] Franek, M.: *Bildentrauschung und Kantenerkennung mit L^p-Regularisierung und Gradientenbeschränkung bei Graustufenbildern.* Diploma thesis; University of Münster 2007

32. [Funk 62] Funk, P.: *Variationsrechnung und ihre Anwendung in Physik und Technik.* Springer; Berlin - Göttingen - Heidelberg 1962 (Grundlehren 94)

33. [Gajewski/Gröger/Zacharias 74] Gajewski, H.; Gröger, K.; Zacharias, K.: *Nichtlineare Operatorgleichungen und Operatordifferentialgleichungen.* Akademie-Verlag; Berlin 1974

34. [Goering/Roos/Tobiska 93] Goering, H.; Roos, H.-G.; Tobiska, L.: *Finite-Element-Methode.* Akademie-Verlag; Berlin 1993, 3rd ed.

35. [Gousseau/Morel 01] Gousseau, Y.; Morel, J.-M.: *Are natural images of bounded variation?* SIAM J. Math. Anal. **33** (2001), 634 – 648

36. [Haber/Modersitzki 07] Haber, E.; Modersitzki, J.: *Intensity gradient based registration and fusion of multi-modal images.* Methods of Information in Medicine **46** (2007), 292 – 299

37. [Hinterberger/Hintermüller/Kunisch/Oehsen/Scherzer 03] Hinterberger, W.; Hintermüller, M.; Kunisch, K.; von Oehsen, M.; Scherzer, O.: *Tube methods for BV regularization.* J. Math. Imaging Vision **19** (2003), 219 – 235

38. [Laird/Wächter 09] Laird, C.; Wächter, A.: *Introduction to IPOPT: A tutorial for downloading, installing, and using IPOPT. Revision No. 1597.* Electronically published: http://www.coin-or.org/Ipopt/documentation/ (accessed at 30. 11. 2009)

39. [Lur'e 75] Lur'e, K. A. (Лурье, К. А.): *Оптимальное управление в задачах математической физики.* Наука; Москва 1975

40. [Maess 88] Maess, G.: *Vorlesungen über numerische Mathematik II.* Akademie-Verlag; Berlin 1988

41. [Meer/Georgescu 01] Meer, P.; Georgescu, B.: *Edge detection with embedded confidence.* IEEE Trans. Pattern Anal. Machine Intell. **23** (2001), 1351 – 1365

42. [Mumford/Shah 89] Mumford, D.; Shah, J.: *Optimal approximations by piecewise smooth functions and associated variational problems.* Comm. Pure Appl. Math. **42** (1989), 577 – 685

43. [Osher/Burger/Goldfarb/Xu/Yin 05] Osher, S.; Burger, M.; Goldfarb, D.; Xu, J.; Yin, W.: *An iterative regularization method for total variation-based image restoration.* Multiscale Model. Simul. **4** (2005), 460 – 489

44. [Perona/Malik 90] Perona, P.; Malik, J.: *Scale-space and edge detection using anisotropic diffusion.* IEEE Trans. Pattern Analysis and Machine Intelligence **12** (1990), 629 – 639

45. [Pickenhain/Wagner 00] Pickenhain, S.; Wagner, M.: *Critical points in relaxed deposit problems.* In: Ioffe, A.; Reich, S.; Shafrir, I. (Eds.): Calculus of Variations and Optimal Control, Technion 98, Vol. II (Research Notes in Mathematics, Vol. 411). Chapman & Hall / CRC Press; Boca Raton etc. 2000, 217 – 236

46. [Rudin/Osher/Fatemi 92] Rudin, L. I.; Osher, S.; Fatemi, E.: *Nonlinear total variation based noise removal algorithms.* Physica D **60** (1992), 259 – 268

47. [Shi/Malik 00] Shi, J.; Malik, J.: *Normalized cuts and image segmentation.* IEEE Trans. Pattern Analysis and Machine Intelligence **22** (2000), 888 – 905

48. [Ting 69a] Ting, T. W.: *Elastic-plastic torsion of convex cylindrical bars.* J. Math. Mech. **19** (1969), 531 – 551

49. [Ting 69b] Ting, T. W.: *Elastic-plastic torsion problem III.* Arch. Rat. Mech. Anal. **34** (1969), 228 – 244

50. [Tröltzsch 05] Tröltzsch, F.: *Optimale Steuerung partieller Differentialgleichungen — Theorie, Verfahren und Anwendungen.* Vieweg; Wiesbaden 2005

51. [Wächter/Biegler 06] Wächter, A.; Biegler, L. T.: *On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming.* Math. Program., Ser. A **106** (2006), 25 – 57

52. [Wagner 96] Wagner, M.: *Erweiterungen des mehrdimensionalen Pontrjaginschen Maximumprinzips auf meßbare und beschränkte sowie distributionelle Steuerungen.* PhD thesis; University of Leipzig 1996

53. [Wagner 99] Wagner, M.: *Erweiterungen eines Satzes von F. Hüseinov über die C^∞-Approximation von Lipschitzfunktionen.* BTU Cottbus, Preprint M-11/1999.

54. [Wagner 03] Wagner, M.: *Variationsrechnung.* Course at the BTU Cottbus in the winter term 2003/04

55. [Wagner 06] Wagner, M.: *Mehrdimensionale Steuerungsprobleme mit quasikonvexen Integranden.* Habilitation thesis; BTU Cottbus 2006

56. [Wagner 08] Wagner, M.: *Quasiconvex relaxation of multidimensional control problems with integrands $f(t, \xi, v)$.* Max-Planck-Institut für Mathematik in den Naturwissenschaften, Leipzig, Preprint Nr. 68/2008. Accepted: ESAIM: Control, Optimisation and Calculus of Variations

57. [Wagner 09] Wagner, M.: *Pontryagin's maximum principle for multidimensional control problems in image processing.* J. Optim. Theory Appl. **140** (2009), 543 – 576

58. [Wagner 10] Wagner, M.: *An optimal control approach to the elastic/hyperelastic image registration problem.* University of Graz, SFB-Report No. 2010–031

59. [Weickert 96] Weickert, J.: *Anisotropic diffusion in image processing.* PhD thesis; University of Kaiserslautern 1996

Last modification: 31. 08. 2010

Addresses / e-mail: *Lucas Franek:* University of Münster, Department of Mathematics and Computer Science, Institute of Computer Science, Einsteinstr. 62, 48149 Münster, Germany. e-mail: `lucas.franek@uni-muenster.de`

Marzena Franek: University of Münster, Department of Mathematics and Computer Science, Institute for Computational and Applied Mathematics, Einsteinstr. 62, 48149 Münster, Germany. e-mail: `marzena.franek@uni-muenster.de`

Helmut Maurer: University of Münster, Department of Mathematics and Computer Science, Institute for Computational and Applied Mathematics, Einsteinstr. 62, 48149 Münster, Germany. e-mail: `maurer@math.uni-muenster.de`

Marcus Wagner: University of Graz, Institute for Mathematics and Scientific Computing, Heinrichstr. 36, A-8010 Graz, Austria. Homepage / e-mail: `www.thecitytocome.de` / `marcus.wagner@uni-graz.at`