Mathematics Meets Medicine – An Optimal Alignment

Bernd Fischer, Eldad Haber, and Jan Modersitzki

Abstract

Image registration is the process of aligning two or more images of the same scene taken at different times, from different viewpoints and/or by different sensors. Image registration is a crucial step in imaging problems where the valuable information is contained in more than one image. Here, spatial alignment is required to properly integrate useful information from the separate images. It is the goal of this note to give an overview on modern techniques in this area. It turns out that the registration problem is an inverse problem which requires regularization and the use of modern optimization methods. We address these issues and supplement it by real-life examples.

1. Introduction

One of the many problems in current image processing is image registration, sometimes also called fusion, matching, or warping. Very often, information obtained from multiple images appears at different poses and is of complementary nature. Therefore spatial alignment is required to properly integrate useful information from separate images. This procedure is called registration. In other words, given a reference and a template image, the goal is to find a transformation, such that the transformed template is similar to the reference image.

There is a vast amount of applications that demand registration including art, astronomy, astrophysics, biology, chemistry, criminology, forensics, genetics, physics, or basically any area involving imaging techniques. More specific applications depending on registration are, for example, remote sensing (constructing a global picture from different partial views), security (comparing current images with a data base), robotics (tracking of objects), and in particular medicine, where compu-
tional anatomy, computer aided diagnosis, fusion of different modalities, intervention and treatment planning, monitoring of diseases, motion correction, radiation therapy, or treatment verification demand for registration. Since imaging techniques, like computer tomography (CT), diffusion tensor imaging (DTI), magnetic resonance imaging (MRI), positron emission tomography (PET), single photon emission computer tomography (SPECT), or ultrasound (US) underwent a remarkable, fascinating, and ongoing improvement in the last decade, a tremendous increase in the utilization of the various modalities in medicine takes place.

Medical image registration algorithms usually estimate the transformation by either following a data driven flow or by minimizing a certain cost function. Flow approaches are very similar to approaches in so-called optical flow in computer vision, a technique used to estimate motion in an image sequence like video. In image registration, flow approaches have been studied for example in \[8, 4, 10\]. For a general introduction to image registration see \[36, 30, 15, 64, 48, 24, 50\] or the overview papers \[11, 71, 62, 46, 23, 42, 20, 30, 37, 59, 65, 19\]. For important theory related contributions see, e.g. \[26, 12\].

This paper focuses on an optimization approach. In the sequel, it is assumed that the images \(T\) and \(R\) are given, smooth, and compactly supported function on a rectangular domain \(\Omega \subset \mathbb{R}^d\), where \(d\) denotes the data dimensionality. Using an appropriate data fitting term \(D\) and a regularizer \(S\) the registration task might be phrased as an optimization problem for the required transformation \(y : \mathbb{R}^d \to \mathbb{R}^d\):

\[
D[T[y], R] + \alpha S[y] + \beta P[y] = \min \text{ st. } y \in M, \tag{1}
\]

where \(\alpha\) is a regularization parameter, \(P\) is a penalty or "soft constrained", \(\beta\) a penalty parameter, and \(M\) a set of admissible transformations.

Optimization based registration can be classified according to the space that the required deformation belongs: rigid or affine linear registration algorithms depend on only a few parameters, while for example spline based approaches may use a very high dimensional transformations space. Rigid registration methods are in general very limited and not descriptive enough and therefore nonrigid approaches are to be used in most applications.

As it is outlined in the next section, image registration is ill-posed which results in ill-conditioning, instability of solutions, and a highly nonconvex cost function. Regularization is introduced to alleviate these issues and also to incorporate user knowledge into the problem formulation. Therefore, image registration is usually phrased as an optimization problem, where the cost function consists of a similarity measure, a regularization, and additional penalty terms that discourage undesirable transformations or constraints that rule out infeasible solutions.

Two major direction can be distinguished. One direction is based on a fixed Tikhonov-regularization, penalizing against a fixed starting point, for example a pre-registration, while the other direction is based on iterative Tikhonov-regularization where the update is regularized. The second approach generally allows for more flexible transformations and thus is an interesting tool for inter-patient registration. The approach is also related to flow like approaches as it can be viewed as a discretization of a time dependent PDE. The first approach is more restrictive and, with a proper choice of the regularization parameter, can be used to ensure a one-to-one transformation, which is very valuable for some applications.

2. The Forward Problem

A major ingredient is the computation of the transformed image \(T[y]\). Since the data is typically given discrete, a first step is interpolation. A common approach is \(d\)-linear interpolation. The advantages of \(d\)-linear interpolation are a continuous representation, a min/max principle, and short computation time. A drawback is that the resulting objective function is non-differentiable. Higher order spline interpolation is an appropriate alternative, resulting in a smoother, differentiable image model. The drawbacks of spline interpolation are under-/overshooting (Gibb’s phenomena), oscillations (introducing local minima in the objective function), and its slightly longer computation time \(60, 61, 64, 57, 25\).

In addition, approximation or scale-space ideas can be used for convexification of the objective function. Rather than interpolating the measurements that are usually be corrupted by noise, a smoother approximation is used \(26, 25\). Image details are omitted and the danger of being trapped into local
minima may be reduced.

Based on the continuous representation, the above so-called Eulerian framework is obtruding, $T[y](x) := T(y(x))$. An alternative, though less commonly used approach is the so-called Lagrangian map, where basically $T^\text{Lagrange}[y](x) := T(y^{-1}(x))$ (in a proper formulation, the inverse transformation is not required). The later approach presents an attractive alternative for some constrained problems since subsets of the domain stay constant during the optimization process and need not to be tracked.

3. Data Fitting

The next step is to quantify image similarity, proximity or distance. Several approaches have been proposed. These measures are based either on image features (like for example markers – tags that are attached to objects before imaging, landmarks – tags deduced after imaging, moments – statistical quantities), substructures (surfaces, level sets, crest lines) or volumetric data. Here, we discuss landmark and $L_2$-norm based distance measures.

Landmarks Based Distances For landmark based approaches, it is assumed that a number of points $r_1, \ldots, r_m \in \mathbb{R}^d$ are specified in the reference image and corresponding points $t_1, \ldots, t_m$ are located in the template image. A commonly used distance measure then reads

$$D^{\text{LM}}[y] = \sum ||y(r^j) - t^j||_{\mathbb{R}^d},$$

(2)

where extensions using different norms are used as well [54].

Several approaches are employed to compute the required transformation. The first one is to restrict the search space to a typically low dimensional parameterizable space $\mathcal{M}$. Common choices for $\mathcal{M}$ are based on rigid, affine linear, or spline transformations. Here, the transformation is a linear combination of some basis functions $q_k$, i.e. $y(x) = \sum q_k(x)w_k$, and it is computed by solving a least squares problem for the coefficients:

minimize $D^{\text{LM}}[y]$ subject to $y \in \mathcal{M}$.  (3)

A second approach is to optimize a certain smoothness $\mathcal{S}$ of the transformation, i.e. minimize $\mathcal{S}[y]$ subject to $D^{\text{LM}}[y] = 0$. This more general concept is discussed further in Section 4. Particularly for landmark based registration, a thin-plate-spline bending energy is used as smoothness measure and the resulting registration is called thin-plate spline registration; see also [13, 3, 43, 54]. Following representer theory [43], it can be shown that the solution to this problem belongs to a certain parameterizable space $\mathcal{M}^{\text{TPS}}$, spanned essentially by translates of a radial basis function related to $\mathcal{S}$ (so-called Thin-Plate-Splines). Hence, this approach is a special case of (3). Numerical solutions can be obtained by solving a linear system of equations for the coefficients.

In practice, in particular in the presence of noise, the exact location of landmarks is a tricky problem. Therefore, the interpolation condition is sometimes replaced by a weighted approximation condition

minimize $D[y] + \theta \mathcal{S}[y]$ subject to $y \in \mathcal{M}^{\text{TPS}}$.

Setting $\theta = 0$ returns (3) whereas $\theta \to \infty$ results in a very smooth transformation which may not fulfill the interpolation constraints; see [54] for a more detailed discussion.

Advantages of landmark based registrations are that their interpretation is easy and intuitive, solutions can be computed fast and efficiently. However, neither the determination of landmarks nor the identification of proper locations of corresponding landmarks is an easy task and a scheme for a fully automatic detection of landmarks in medical images is still missing.

Volumetric Distances To overcome these drawbacks one is bound to consider the whole images. Probably the most intuitive volumetric distance measure is the so-called sum of square difference (SSD) (also $L_2$-norm of the image difference or energy of image distance),

$$D^{\text{SSD}}[y] = ||T[y] - R||_{L_2(\Omega)} = \int_\Omega (T[y] - R)^2 \, dx.$$  (4)

It has been proved to be robust and very effective for images of one modality. It is based on the assumption that there exists a transformation $y$ with $T[y](x) = R(x)$, which most often does not hold. For example, fusion of different modalities is a typical registration task which contradicts this assumption. Here, cross-correlation (and variants), normal-
ized gradient fields \cite{29}, and in particular mutual information \cite{63, 10} are commonly used alternatives; see, e.g., \cite{52, 35, 48} for a detailed overview.

4. Regularization

Image registration is inherently ill-posed. For every spatial location $x$, one asks for a vector valued transformation $y(x)$, but in general only scalar information is provided. Regularization is thus important and inevitable. It should be mentioned that the focus of regularization in image registration is on the existence rather than on the uniqueness of the underlying optimization problem \cite{1}. Regularization is achieved by selecting an admissible set $M$ and/or by choosing an application dependent Sobolev seminorm (bi-linear form)

$$S[y] = ||B[y]||_{L_2(\Omega)}^2 = \int_\Omega \langle B[y], B[y] \rangle \, dx,$$

where in particular the following choices are common (omitting some parameters)

$$B_{\text{diff}}[y] = (\nabla y_1, \ldots, \nabla y_d) \text{ (diffusion)},$$

$$B_{\text{elas}}[y] = (\nabla y_1, \ldots, \nabla y_d, \nabla \cdot y) \text{ (elastic)},$$

$$B_{\text{curv}}[y] = (\Delta y_1, \ldots, \Delta y_d) \text{ (curvature)};$$

see \cite{6, 16, 18, 48, 31}. Unlike other ill-posed problems, the particular regularizer and in particular the choice of boundary conditions can affect the solution considerably. As pointed out for the above example, the influence becomes less prominent, if the images have more structure.

Another important issue is the choice of the regularization parameter. Typically, this parameter is hand-tuned. For an automatic choice of the regularization parameter, a continuation method has been suggested in \cite{27}.

5. Soft Constraints and Penalties

Another step towards a more reliable outcome of the registration is to add a penalty $P$ penalizing unwanted solutions. In contrast to a regularizer, which is needed to guarantee solutions of the registration problem, the penalty is an add-on which may serve as an instrument for incorporating user knowledge. Examples for penalty terms include the deviation from user supplied landmarks \cite{39}, volume preservation \cite{53} or local rigidity \cite{11, 15, 54, 49}.

6. Constraints

In contrast to the penalty approach, where unwanted solutions are penalized, hard constraints are used to rule out these transformations. The feasible set is either described explicitly or in terms of constraints $C$. Explicit descriptions are, for example, parameterized transformations like rigid, affine linear, or diffeomorphic spline transformations \cite{17, 2, 55}. Implicit constraints range from equality constraints $C[y] = 0$ like landmarks constraints \cite{17} and space segregation \cite{25} (low dimensional, linear), or local rigidity (high dimensional, almost linear), volume preservation \cite{25} (high dimensional, non-linear) to inequality constraints. The displayed equality constraints match the penalty terms of the previous section. However, instead of penalizing unwanted solutions, the constraints completely rule them out. Moreover, by using hard constraints one gets rid of the additional penalty parameter.

7. Optimization

The backbone of the solution schemes are numerical optimization techniques, see, e.g., \cite{21, 22, 11, 51}.

In the literature one may find essentially three different optimization approaches within a registration procedure. The first one is based on practitioners heuristics. The second class of approaches aims for the necessary condition for a minimizer of the objective function in \cite{1}, which turns out to be a nonlinear system of partial differential equations. Subsequently, the partial differential equation, equipped with appropriate boundary conditions, is solved by some sort of discretization technique; see also \cite{8, 9, 58, 16, 17, 34, 48, 31}. The third class interchanges the role of optimization and discretization. It is based on a sequence of nested discretizations of \cite{1} \cite{27}. Each discretization leads to a finite dimensional optimization problem. Starting with a coarse discretization, a minimizer is computed which then serves as a starting point for the optimization problem associated with the finer discretization. If the coarse grid solution is a good starting guess, the optimization on the finer discretization is more like a correction step and is expected to be performed with low computational costs. To arrive at a fast converging scheme, care has to be taken dur-
ing modeling. If the objective function is sufficiently smooth, Newton-type optimization techniques can be applied. A major advantage, as compared to the PDE-based approaches, is that sound globalization techniques and stopping criteria are available. Focusing on optimization aspects rather than smoothing and grid transfer operations constitutes a difference to similar multigrid type techniques; see, e.g., [32, 33].

No matter which strategy is employed, there is always the danger of being trapped by an unwanted minimum. To reduce this problem, four main ideas are used. These ideas may be seen as convexification of the overall optimization problem. The first is to start with a pre-registration that generates an acceptable starting guess in a neighborhood of a minimizer. The second step is to add a regularization $S$ and/or penalties $P$, penalizing unwanted solutions. The third idea is to solve the problem in a multi-layered fashion, i.e. on a sequence of discretizations ranging from coarse to fine, where on a coarse level only main characteristics of the underlying images are visible and are resolved and where subsequently on finer levels more and more characteristics are added and resolved. Here, the point is to use a multi-resolution of the images. The fourth concept is along the same lines, details are suppressed for the start-up phase and then are gradually added. This way one arrives at a multi-scale technique which of course can be mixed with the multi-resolution approach.

REFERENCES


Mixed Integer Linear Sets

Kent Andersen
Department of Mathematical Sciences, University of Copenhagen, Denmark (kha@math.ku.dk)

Robert Weismantel
Faculty of Mathematics, Otto-von-Guericke-Universität Magdeburg, Universitätsplatz 2, 39106 Magdeburg, Germany (weismantel@imo.math.uni-magdeburg.de)

1. Introduction

From a practical perspective, mixed integer optimization represents a very powerful modeling paradigm. Its modeling power, however, comes with a price. The presence of both integer and continuous variables results in a significant increase in complexity with respect to geometric, algebraic, combinatorial and algorithmic properties. Specifically, the theory of cutting planes for mixed integer linear optimization is not yet at a similar level of development as in the pure integer case. The goal of this survey is to examine a new geometric approach based on lattice point free polyhedra and use it to develop a cutting plane theory for mixed integer sets. We expect that these novel developments will naturally extend the theory that has been developed for the pure integer case and shed some light on the additional complexity that goes along with mixing discrete and continuous variables. The point of departure for studying mixed integer linear sets is the fact that for a polyhedron $P \subseteq \mathbb{R}^{n+d}$ with feasible mixed integer set $Q = P \cap (\mathbb{Z}^n \times \mathbb{R}^d)$, the set $\text{conv}(Q)$ is a polyhedron. Hence,

$$c^* = \max \{ c^T x + g^T y \mid (x, y) \in Q \} = \max \{ c^T x + g^T y \mid (x, y) \in \text{conv}(Q) \}.$$ 

The following questions emerge:

(1) Which geometric tools are needed in order to understand $\text{conv}(Q)$?

(2) Which algebraic tools are needed to generate cutting planes?

(3) Can (1) and (2) be turned into a finite algorithm for computing $c^*$?

Whereas in the pure integer case geometric principles such as the rounding off hyperplanes [10] and the lift-and-project approach [5] lead to finite cutting plane procedures [12], these questions remain challenging in the presence of both discrete and continuous variables. Indeed, already for a mixed integer program in two integer and one continuous variable, it requires major efforts to design a finite cutting plane algorithm.

Example 1 [11] For the mixed integer program in variables $x_1, x_2 \in \mathbb{Z}_+$ and $y \in \mathbb{R}_+$,

$$\begin{align*}
\max & \quad + y \\
- x_1 & \quad + y \leq 0 \\
-x_2 + y & \leq 0 \\
+ x_1 & \quad + x_2 + y \leq 2
\end{align*}$$

the optimal LP relaxation is attained at the point $(\frac{5}{2}, \frac{5}{2})$. The inequality $y \leq 0$ is valid, because the triangle $x_1 \geq y, x_2 \geq y, x_1 + x_2 \leq 2 - y$ contains an integer point if and only if $y = 0$. This inequality, however, cannot be derived using a finite number of rounds of rounding hyperplanes.

2. Mixed Integer Rounding and Beyond

Major developments with regard to a cutting plane theory for mixed integer linear optimization problems are based on the following principles that we briefly outline below. For an introduction to the theory of linear integer and mixed integer optimization we refer to [7, 23, and 24].
Mixed integer rounding

This line of research began with the developments of Ralph Gomory [14]. The material is well summarized in [23]. Gomory’s integer and mixed integer cut can be derived from the simplex tableau associated with a given vertex of the continuous relaxation of a given mixed integer program. Whereas the derivation in the pure integer case is elementary as it is only based on rounding up and down fractional coefficients, its mixed integer counterpart is much more involved. This follows from the fact that the fractionalities of the coefficients associated with the continuous variables must be taken into consideration. Interestingly, Gomory’s fractional cut is a special case of a so-called mixed integer rounding cut introduced by Nemhauser and Wolsey [23]. The mixed integer rounding cut in its basic form is the only non-trivial inequality that is required to describe the convex hull of the two dimensional mixed integer set

\[ \{(x, y) \in \mathbb{Z} \times \mathbb{R}_+ \mid x + y \geq b\}, \]

where \( b \in \mathbb{Q} \) as illustrated below. In formula, the cut reads

\[ x + \frac{1}{1 - f(b)} y \geq \lceil b \rceil, \]

where \( f(b) \) denotes the fractional part of \( b \in \mathbb{Q} \). The principle of mixed integer rounding can be naturally applied to general models by aggregating variables. Following this scheme one may derive the so-called Gomory-fractional cut, see [23] for further details. In certain cases mixed integer rounding cuts can be manipulated further and combined to generate additional cutting planes. This scheme is known as mixing and was introduced by Pochet and Günlük [18].

Modular Arithmetic and Superadditive Functions

Rounding and mixed integer rounding can be viewed as evaluating a special superadditive function that is applied to the constraint matrix associated with a mixed integer program. In a similar vein, cyclic group relaxations of a mixed integer program such as the corner polyhedron can be described by superadditive functions. This line of research has been initiated by Gomory [13] and further extended and refined by Ellis Johnson [17] [16]. Our survey is not further linked to this kind of approach. We refer the interested reader to volume 96 of Mathematical Programming B, 2003 for recent survey papers on the subject.

Cuts From Several Tableau Rows

In [3] the geometry of the integer points in a translate of a cone rooted at a rational point has been investigated. This investigation led to some unexpected links between lattice point free bodies and the derivation of inequalities for a mixed integer set described by two rows of a simplex tableau. From the theory of linear programming it follows that a vertex \( x^* \) of the LP corresponds to a basic feasible solution of a simplex tableau associated with subsets \( B \) and \( N \) of basic and nonbasic variables

\[ x_i + \sum_{j \in N} \bar{a}_{i,j} x_j = \bar{b}_i \text{ for } i \in B. \]

Any row associated with an index \( i \in B \) such that the corresponding variable \( x_i \) is required to be integer and \( \bar{b}_i \notin \mathbb{Z} \) gives rise to a relaxation

\[ X(i) := \{ x \in \mathbb{R}_+^{\lceil N \rceil} \mid \bar{b}_i - \sum_{j \in N} \bar{a}_{i,j} x_j \in \mathbb{Z} \} \]

that can be used to generate inequalities that are violated by \( x^* \). Indeed, Gomory’s mixed integer cuts [14] and mixed integer rounding cuts are derived from such a basic set \( X(i) \) using additional information about integrality of some of the variables. Let
us next consider an index set \( \tilde{I} \subseteq B \cap I \) of cardinality greater than or equal to two. With this index set of basic variables one associates the mixed integer set \( X(\tilde{I}) \) defined as all the nonnegative points in \( \mathbb{R}^{|N|} \) that satisfy

\[
\tilde{b}_i - \sum_{j \in N} \tilde{a}_{i,j} x_j \in \mathbb{Z} \text{ for } i \in \tilde{I}.
\]

This set defines geometrically a translate of a polyhedral cone rooted at a given vertex. When \( \tilde{I} \) has cardinality two, then a description of \( \text{conv}(X(\tilde{I})) \) is known [3]. This description is given in terms of the maximal lattice point free polyhedra that describe the projection of the facets to the space of integer variables. For further results in this direction we refer to [8, 9].

For subsets \( \tilde{I} \) of cardinality greater than two, the projection of any facet to the space of integer variables is still a maximal lattice point free polyhedron (except in some minor special cases). However, a classification of the lattice point free polyhedra is unknown in higher dimensions. In fact, this classification is an important element in the proof given in [3]. This motivates us to study the effect of adding specialized cuts from three-row relaxations. One step into this direction is to numerically evaluate the effect of adding families of cutting planes from three rows of a simplex tableau compared to two-row relaxations. At this point one may expect that novel results about the approximation of disjunctions derived from approximations of high dimensional maximal lattice point free polyhedra will help to cope with the latter question, see also [6].

### 3. Split Cuts and Disjunctions

Disjunctive programming originates from the work of Egon Balas [4]. A disjunction of a mixed integer set requires to partition the feasible set into two or more parts in a way so that at least one element of the partition contains no feasible mixed integer point. Then one can convexify the non-empty elements of the partitioning and arrive at an improved formulation. The cuts obtained from the convexification step are called disjunctive cuts.

In the mixed 0 – 1-case, disjunctive cuts can be derived elegantly from the projection of a higher dimensional linear reformulation of the original problem (based on the identity \( z^2 = z \) for a binary variable \( z \)). This scheme is known under the name “Lift-and-project”, see [22, 25, 5]. When applied iteratively, the convex hull of the mixed 0 – 1-problem is obtained.

Split cuts were introduced by Cook, Kannan and Schrijver in 1990 [11]. They are a special case of disjunctive cuts in the sense that they are based on specialized partitions of a mixed integer feasible region \( F \) into three parts, \( F_0 = F \cap \{ x \mid \pi^T x \leq \pi_0 \} \), \( F_1 = F \cap \{ x \mid \pi^T x \geq \pi_0 + 1 \} \) and \( F_2 = F \cap \{ x \mid \pi_0 < \pi^T x < \pi_0 + 1 \} \). If \( \pi, \pi_0 \) are integral and \( \pi_i = 0 \) for all indices \( i \) corresponding to the continuous variables, then \( F_2 \) cannot contain any mixed integer point. That is, a split cut comes from the convexification \( \text{conv}\{ F_0, F_1 \} \), see the figure below for illustration.

Interestingly, both Gomory’s fractional cuts and the mixed integer rounding cuts are split cuts.

It is an important fact that a cutting plane algorithm based on split cuts cannot always guarantee finite convergence in the mixed integer case [11]. This is a quite negative result as it limits at least theoretically the applicability of split cuts to general mixed integer linear programs. On the other hand, this result inspires a number of interesting research questions that are at least partially discussed here. The major research challenge in this context is to provide insight into the geometry of the disjunction that is to be used in order to guarantee that a corresponding disjunctive cutting plane algorithm can solve a given mixed integer program in finitely many rounds. Indeed, it was recently shown in [19] that if one allows for arbitrary disjunctions, then finite termination of a cutting plane algorithm can be
guaranteed.

4. Splits and Lattice Point Free Polyhedra

In [11] it is shown that applying split cuts iteratively does not suffice to generate the cut \( y \leq 0 \) that we developed in Example 1. This demonstrates that more complicated disjunctions must be taken into consideration so as to result in a finite cutting plane algorithm.

The basic idea of generalizing the operation of adding split cuts is to view split disjunctions as a special family of lattice point free polyhedra. For a survey on lattice-point-free convex sets we refer to [21].

Let \( L \subseteq \mathbb{R}^n \) be a convex set. By \( \text{int}(L) \) we denote the interior of \( L \). We call \( L \) lattice-point-free if \( \text{int}(L) \cap \mathbb{Z}^n = \emptyset \). We call lattice point free closed convex sets that are maximal with respect to inclusion split bodies. It turns out that split bodies are full dimensional rational polyhedra. In fact, a split body can always be written as the sum of a polytope \( P \) and a linear space \( \mathcal{L} \). This fact suggests to associate a split dimension with every split body. The split dimension of a split body \( L \) is denoted by \( \text{dims}(L) \) and defined to be the dimension of the corresponding polytope \( P \). For instance, split cuts arise from split bodies of dimension one: a split set can be written as the sum of a one-dimensional line segment \( P := \text{conv}(\{v^1, v^2\}) \) and a \((n-1)\) dimensional linear space \( \mathcal{L} := \{x \in \mathbb{R}^n : \pi^Tx = 0\} \), where \( \pi^Tv^1 = \pi_0 \) and \( \pi^Tv^2 = \pi_0 + 1 \). Besides the split dimension of a split body, there is a second important criterion to distinguish such sets. This leads us to the notion of the maximum facet width. Recall that for \( L \subseteq \mathbb{R}^n \) and \( \pi \in \mathbb{Z}^n \) the width of \( L \) along \( \pi \) is the number \( w_\pi(L) = \max\{\pi^T x \mid x \in L\} - \min\{\pi^T x \mid x \in L\} \). For a split body \( L = \{x \in \mathbb{R}^n \mid Ax \leq b\} \) with \( A \in \mathbb{Z}^{m \times n} \), the maximum facet width \( f(L) \) is the number \( f(L) = \max\{w_{\pi_i}(L) \mid i = 1, \ldots, m\} \). Notice that a split body \( L \) has maximum facet width equal to one if and only if \( L \) is a split body of split dimension one. For instance, the split body \( \text{conv}\{(0,0),(2,0),(0,2)\} \) of split dimension two has maximum facet width equal to two.

We define as

\[
cuts_P(L) = \text{conv}\{(x,y) \in P \mid x \notin \text{int}(L)\},
\]

the operation of adding cuts to \( P \) from the lattice-point-free polyhedron \( L \). This is illustrated in the two figures below. The first important remark is that these new inequalities are valid for any mixed-integer problem that has \( P \) as its continuous relaxation. Observe that the outer description of \( \text{cuts}_P(L) \) may involve several new inequalities.
5. The Split Body Closure

We are now prepared to introduce the closure operation based on split bodies. In doing so, we assume that one is given a polyhedron $P \subseteq \mathbb{R}^{n+m}$. The corresponding mixed integer feasible region is defined to be $P \cap (\mathbb{Z}^n \times \mathbb{R}^m)$. Let the parameter $\omega$ be an upper bound on the maximum facet width of the split bodies. Let $\mathcal{F}_\omega$ denote the set of all split bodies in $\mathbb{R}^{n+m}$ of maximum facet width less or equal than $\omega$. The $\omega$-closure of $P$ is defined as

$$sc_\omega(P) := \bigcap_{L \in \mathcal{F}_\omega} \text{cuts}_P(L).$$

The set $sc_\omega(P)$ is convex, by definition. Note however, that here the intersection is taken over an infinite number of bodies. This is why there is no obvious reason to believe that $sc_\omega(P)$ is always a polyhedron.

In fact, it has been shown in [11] and [1] that the conjecture is true when $\omega = 1$. In this special case, the result states that the usual split closure of a polyhedron is again a polyhedron.

It has been shown recently in [2] that for all $\omega$, the set $sc_\omega(P)$ is a polyhedron. The proof of this result is quite involved and it would be beyond the scope of this survey to present details of it. Let us instead shed some light on a number of questions for which a positive answer is conceivable, given that $sc_\omega(P)$ is a polyhedron.

In particular, in [2] it is shown that with every mixed integer linear program there is associated an $\omega$ with the following property: Any cutting plane algorithm using split bodies of maximum facet width strictly less than $\omega$ cannot solve any optimization problem $\{\max c^T x, \ x \in P_I\}$ in finitely many rounds. This result generalizes the phenomenon of non termination if one uses regular split cuts only and applies these to Example 1.

Another issue concerns cutting plane proofs based on split bodies. Since $sc_\omega(P)$ is a polyhedron, every valid inequality for a mixed integer linear program can be derived from recursively applying the operation $\text{cuts}_P(L)$ working with a family $\mathcal{F}$ of split bodies $L$. The maximal maximum facet width of such a body $L$ in the family serves as a measure of the “complexity” for deriving the inequality under investigation. In Example 1 the split body proof of the inequality $y \leq 0$ must have maximal maximum facet width equal to two.

6. A Perspective

The concept of split bodies that we introduced in this paper motivates numerous research questions. Below we exhibit three such directions.

Cutting Plane Methods Based on Split Bodies

As we already pointed out it was shown by Jörg [19] that if one allows for arbitrary disjunctions, then a finite cutting plane algorithm for mixed integer optimization can be developed. It remains, however, open to understand precisely the connection between the geometry of the split bodies and the event of finite termination.

The authors conjecture that the following cutting plane algorithm will work: One repeatedly computes the optimal face of a linear relaxation (not just a vertex). If this face of the relaxation does not contain any mixed integer point, then using a mixed integer Farkas type lemma one ought to be able to generate a split body of split dimension equal to the dimension of the face of the relaxation that can be used to set up a disjunction. This disjunction is then used for generating cutting planes. Is it conceivable that a repeated application of this strategy will terminate and return a mixed integer optimal solution after a finite number of rounds? If this statement turns out to be correct, then the complexity of the multi-term disjunction is precisely equal to the dimension of the optimal face of the relaxation. This would provide us with a first link between the geometry of the split bodies and the face-complex of the mixed integer program.

The Tailing Off Effect

A second important topic is to understand the effect of suppressing “tailing off” by increasing the split dimension for cutting plane generation. More precisely, suppose that one starts with cuts arising from one dimensional split bodies. Usually, after a number of rounds the improvement of the objective function becomes small (tailing off). Once the improvement in terms of the objective function is below
a certain threshold, disjunctions based on two- and even higher dimensional split bodies are to be applied. Which conditions must the split body satisfy in order to avoid this tailing off effect? An answer to this question will allow us to design a hierarchical scheme for using more and more complex split bodies. Indeed, the development of such a hierarchical cutting plane algorithm will lead to a geometrically insightful extension of all the traditional finite cutting plane algorithms for the pure integer case.

**Approximation of split bodies**

Split bodies are lattice point free in their interior and maximal with respect to this property. In dimension two such bodies are either triangles or quadrangles. This nice structure does not carry over to higher dimensions. In higher dimensions a characterization of all the lattice point free polyhedra is not available. This in turn implies that from an algorithmic point of view there is little hope to be able to compute the closure $\text{sc}_\omega(P)$. This fact suggests to study approximations of $\text{sc}_\omega(P)$ that result from adding only those cuts that one obtains from the disjunctions associated with a finite and well selected family of split bodies. The analysis of such an approximation is quite nicely linked to the concept of the strength of the cutting planes that one obtains from a disjunction. The notion of strength has been introduced by Goemans in [15]. We believe that the strength of the cuts from one additional disjunction ought to be linked to the size of the subregion of the new disjunction which is not covered by the already available disjunctions. If results along these lines are true, then the road is paved to analytically describe the effect of adding, incrementally, cutting planes from $d$-dimensional splits and compare this new relaxation with the previous one. This way one can evaluate the gain from adding all cuts that one can derive from a specialized family of split bodies.

**REFERENCES**


---

**Bulletin**

Special Issues Dedicated to Henry Wolkowicz

Call for Papers

Special Issue of Mathematical Programming (Series B) on Cone Programming and its Applications

Dedicated to Henry Wolkowicz on the occasion of his 60th birthday

We invite submissions of novel research articles for a forthcoming issue of Mathematical Programming (Series B) on cone programming and its applications, in the broadest sense.

This special issue is associated with the Modeling and Optimization: Theory and Applications (MOPTA) 2008 conference. The issue will be dedicated to Henry Wolkowicz on the occasion of his 60th birthday. Professor Wolkowicz is known worldwide as a major contributor to the theory, algorithms, and applications of cone programming, including semidefinite programming.

We invite papers that address the following topics, individually or in combination:

- Theory of moments and positive polynomials;
- Practical interior-point methods for large-scale cone programming;
- High-impact applications to areas such as combinatorial optimization and global optimization.

The deadline for submission of full papers is December 31, 2008. We aim at completing a first review of all papers by April 30, 2009.

Electronic submissions to the guest editors in the form of pdf files are encouraged. All submissions will be refereed according to the usual standards of Mathematical Programming. Information about Mathematical Programming (Series B) including author guidelines and other special issues in progress, is available at
Mathematical Programming Computation: Call for Papers

The Mathematical Programming Society will publish the new journal *Mathematical Programming Computation (MPC)* beginning in 2009. The journal is devoted to computational issues in mathematical programming, including innovative software, comparative tests, modeling environments, libraries of data, and/or applications. A main feature of the journal is the inclusion of accompanying software and data with submitted manuscripts. The journal’s review process includes the evaluation and testing of the accompanying software. Where possible, the review will aim for verification of reported computational results.

The full contents of the journal will be made freely available on the society-run web site [mpc.zib.de](http://mpc.zib.de). MPC will also be published together with Springer Verlag, in both print and online versions; MPS members will receive the print version of the journal as part of their membership benefits.

MPC supports the creation and distribution of software and data that foster further computational research. The opinion of the reviewers concerning this aspect of the provided material is a considerable factor in the editorial decision process. Another factor is the extent to which the reviewers are able to verify the reported computational results. To these aims, authors are highly encouraged to provide the source code of their software. Articles describing software where no source code is made available are acceptable, provided reviewers are given access to executable codes that can be used to evaluate reported computational results.

Beginning November 1, 2008, articles can be submitted in Adobe PDF format through MPC’s web-based system at [mpc.zib.de](http://mpc.zib.de). Software and supplementary material can also be submitted through this system. Prior to November 1, submissions can be sent to William Cook at mpc@isye.gatech.edu.

Further information can be found on the web page [www.isye.gatech.edu/~wcook/mpc/](http://www.isye.gatech.edu/~wcook/mpc/)

William Cook, Georgia Institute of Technology
Thorsten Koch, Konrad-Zuse-Zentrum Berlin

---

**Chairman’s Column**

The Optimization meeting in Boston in 2008 was an overwhelming success and many thanks are due to Kurt and Sven for their efforts in this regard. The expected attendance had been in the range of 400-450. Actual attendance turned out to be 543 paying attendees. The meeting featured 530 presentations. There were 80 minisymposia, which typically have four speakers per session, and 179 individual contributed presentations. I hope that you enjoyed this meeting as much as I did. I am looking forward to the next meeting in 2011 and continuing to explore options for its location. Please send any suggestions to help our planning to our program director (Steve Vavasis) or to me.

At the end of 2007 (the last complete membership year), the activity group showed a membership of 920, its highest ever. Of these, 397 were students. As of October 31, 2008 the number of members had grown to 1,103 for 2008; of these, 520 were students (reflecting the fact that student members can now choose 2 free activity groups). If you have not yet renewed your membership, please do so at your earliest convenience: [https://my.siam.org/forms/join_siag.htm](https://my.siam.org/forms/join_siag.htm).

While avoiding the inevitable stack of "linear programming" exams to mark, and looking out of my office window as the snow piles accumulate, I am heartened by the opportunities that await us all in the coming year. I hope to see all of you either at the SIAM Annual Meeting in Denver on July 6-10, or at the International Symposium on Mathematical Programming that is returning to Chicago on August 23-28 for the 60th Anniversary of the zeroth ISMP (who said we were all Fortran programmers!).
2009 has all the makings of another great year for Optimization.

**Michael C. Ferris**, SIAG/OPT Chair
Computer Sciences Department
University of Wisconsin-Madison
1210 West Dayton Street,
Madison, WI 53706
USA
ferris@cs.wisc.edu
http://pages.cs.wisc.edu/~ferris/

---

**Comments from the Editor**

This is my first issue of View-and-News, and I am excited and honored to be its new editor. My aim is to broaden the appeal of View-and-News, and, with your help, increase the number of issues per year. Please contact me with suggestions for articles, or special issues!

The present issue presents two articles that highlight the increasing diversity of our activity group. Both articles describe work presented at the SIAM Optimization Conference in Boston. The first paper by Bernd Fischer, Eldad Haber, and Jan Modersitzki provides an overview of a novel approach to image registration. The aim of image registration is to align images to integrate additional information. The second paper by Kent Andersen and Robert Weismantel presents a geometric approach to generalize the theory of cutting planes from pure integer optimization to mixed-integer optimization. It contains both a review of some recent developments, and, interestingly, a number of conjectures and open problems.

I am taking over from a very talented and accomplished editor, Luís Vicente. Luís edited View-and-News for four years (2003–2008), producing eight interesting and entertaining high-quality issues. He set a standard that will be hard to match. Our activity group owes him a debt of gratitude: *Luís, estamos muito gratos por seu fantástico trabalho editorial!*

**Sven Leyffer**, Editor
Mathematics and Computer Science Division
Argonne National Laboratory
Argonne, IL 60439, USA
leyffer@mcs.anl.gov
http://www.mcs.anl.gov/~leyffer/