Linear and Mixed Integer Optimization: The Solution Methods — Just a Glimpse

Martin Grötschel

CO@W Berlin

28.09.2015

Contents

Contents

Typical optimization problems

$$\max\ f(x)\ \ or\ \min\ f(x)$$
$$g_i(x) = 0,\quad i = 1, 2, \ldots, k$$
$$h_j(x) \le 0,\quad j = 1, 2, \ldots, m$$
$$x \in \mathbf{R}^n\ (\text{and}\ \ x \in S)$$

„general“
(nonlinear)
program
NLP

$$\min c^T x$$
$$Ax = a$$
$$Bx \le b$$
$$x \ge 0$$
$$(x \in \mathbf{R}^n)$$
$$(x \in \mathrm{k}^n)$$
$$(x \in \mathbb{Q}^n)$$

linear
program
LP

$$\min c^T x$$
$$Ax = a$$
$$Bx \le b$$
$$x \ge 0$$
$$\text{some } x_j \in \mathbf{Z}$$
$$(x \in \{0,1\}^n)$$

(linear)
0/1-
mixed-
integer
program
IP, MIP

All data are rational.

program = optimization problem

Linear Programming: a very brief history

1826/1827 Jean Baptiste Joseph Fourier (1786-1830): rudimentory form of simplex method in 3 dimensions.

1939 L. V. Kantorovitch (1912-1986): Foundations of linear programming (Nobel Prize 1975)

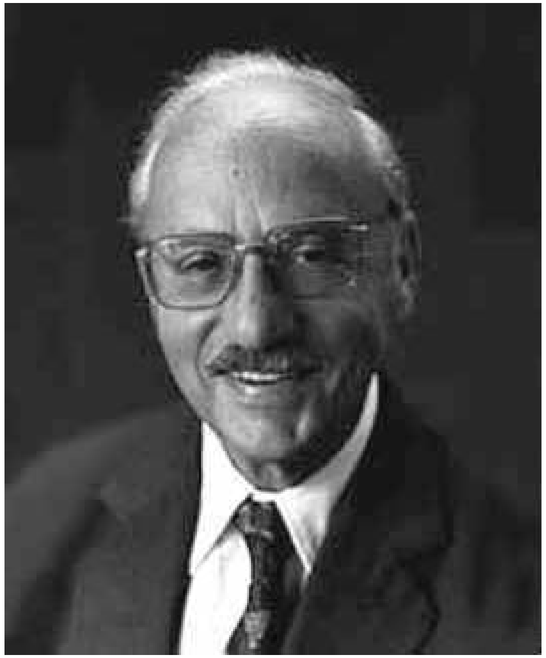
1947 G. B. Dantzig (1914-2005): Primal simplex algorithm

1954 C.E. Lemke: Dual simplex algorithm

1953 G.B. Dantzig,
1954 W. Orchard Hays, and
1954 G. B. Dantzig & W. Orchard Hays: Revised simplex algorithm

1979 L. G. Khachiyan (1952-2005): The ellipsoid method

1984 N. Karmarkar: Interior point methods



Optimal use of scarce ressources foundation
and economic interpretation of LP

Leonid V. Kantorovich Tjalling C. Koopmans

Nobel Prize for Economics 1975

The Decade of the 70's: Practice

Interest in optimization flowered

- Large scale planning applications particularly popular, significant difficulties emerged
- Building applications was very expensive and very risky
- Technology just wasn't ready:
  - LP was slow and
  - Mixed Integer Programming was impossible.

OR could not really "deliver" – with some exceptions, of course

The ellipsoid method of 1979 was no practical success.

The Decade of the 80's and beyond

Mid 80's:

- There was perception was that LP software had progressed about as far as it could.

There were several key developments

- IBM PC introduced in 1981
  - Brought personal computing to business
- Relational databases developed. ERP systems introduced.
- 1984, major theoretical breakthrough in LP
  N. Karmarkar, "A new polynomial-time algorithm for linear programming", Combinatorica 4 (1984) 373-395
  (Interior Point Methods, front page New York Times)

The last ~30 years: Remarkable progress

- We now have three competitive algorithms:
  Primal & Dual Simplex, Barrier (interior points)

Contents

Application of LP & MIP - I

Transportation-airlines

- Fleet assignment
- Crew scheduling
- Ground personnel scheduling
- Yield management
- Fuel allocation
- Passenger mix
- Booking control
- Maintenance scheduling
- Load balancing/freight packing
- Airport traffic planning
- Gate scheduling/assignment
- Upset recover and management

Transportation-other

- Vehicle routing
- Freight vehicle scheduling and assignment
- Depot/warehouse location
- Freight vehicle packing
- Public transportation system operation
- Rental car fleet management

Process industries

- Plant production scheduling and logistics
- Capacity expansion planning
- Pipeline transportation planning
- Gasoline and chemical blending

Application of LP & MIP - II

Financial

- Portfolio selection and optimization
- Cash management
- Synthetic option development
- Lease analysis
- Capital budgeting and rationing
- Bank financial planning
- Accounting allocations
- Securities industry surveillance
- Audit staff planning
- Assets/liabilities management
- Unit costing
- Financial valuation
- Bank shift scheduling
- Consumer credit delinquency management
- Check clearing systems
- Municipal bond bidding
- Stock exchange operations
- Debt financing

Manufacturing

- Product mix planning
- Blending
- Manufacturing scheduling
- Inventory management
- Job scheduling
- Personnel scheduling
- Maintenance scheduling and planning
- Steel production scheduling

Coal Industry

- Coal sourcing/transportation logistics
- Coal blending
- Mining operations management

Forestry

- Forest land management
- Forest valuation models
- Planting and harvesting models

Application of LP & MIP - III

Agriculture

- Production planning
- Farm land management
- Agricultural pricing models
- Crop and product mix decision models
- Product distribution

Public utilities and natural resources

- Electric power distribution
- Power generator scheduling
- Power tariff rate determination
- Natural gas distribution planning
- Natural gas pipeline transportation
- Water resource management
- Alternative water supply evaluation
- Water reservoir management
- Public water transportation models
- Mining excavation models

Oil and gas exploration and production

- Oil and gas production scheduling
- Natural gas transportation scheduling

Communications and computing

- Circuit board (VLSI) layout
- Logical circuit design
- Magnetic field design
- Complex computer graphics
- Curve fitting
- Virtual reality systems
- Computer system capacity planning
- Office automation
- Multiprocessor scheduling
- Telecommunications scheduling
- Telephone operator scheduling
- Telemarketing site selection

Application of LP & MIP - IV

Food processing
- Food blending
- Recipe optimization
- Food transportation logistics
- Food manufacturing logistics and scheduling

Health care
- Hospital staff scheduling
- Hospital layout
- Health cost reimbursement
- Ambulance scheduling
- Radiation exposure models

Pulp and paper industry
- Inventory planning
- Trim loss minimization
- Waste water recycling
- Transportation planning

Textile industry
- Pattern layout and cutting optimization
- Production scheduling

Government and military
- Post office scheduling and planning
- Military logistics
- Target assignment
- Missile detection
- Manpower deployment

Miscellaneous applications
- Advertising mix/media scheduling
- Pollution control models
- Sales region definition
- Sales force deployment

Contents

Linear Program (LP)

Characteristics

Objective

- linear function

Feasible region

- described by linear constraints

Variable domains

- real values

$$\min c^T x$$

$$Ax = a$$

$$Bx \le b$$

$$x \ge 0$$

Observation

1. Optimization algorithms for linear programs can solve the feasibility/membership problem for the associated polyhedron.
2. With a membership algorithm for a polyhedron one can solve any linear optimization problem over the polyhedron. (binary search or combining the primal and the dual program)
3. An LP "min/max $c^T x$, $Ax=b$, x□0" is often called "standard form. Note, though, that each of the methods to be discussed has a slightly different "standard form". They are all "trivially equivalent" in the sense that one form can be easily transformed into the other and there is a simple correspondence between feasible and optimal solutions.

Algorithms for the solution of linear programs

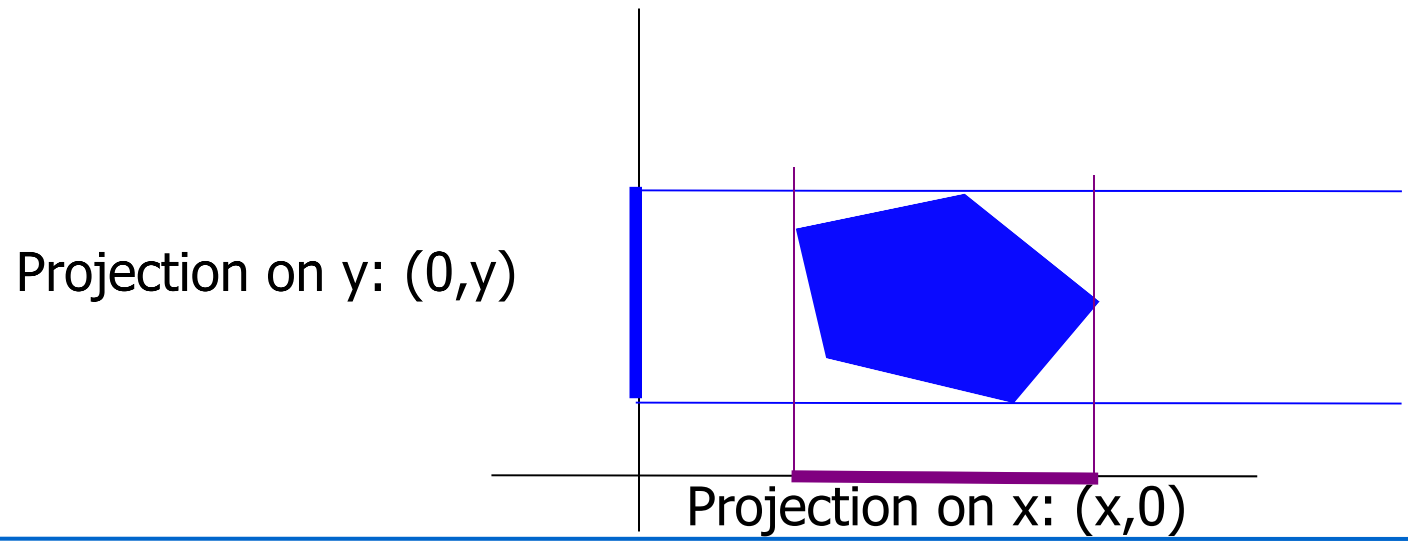
1. Fourier-Motzkin Elimination
2. The Primal Simplex Method
3. The Dual Simplex Method
4. The Ellipsoid Method
5. Interior-Point/Barrier Methods
6. Lagrangian Relaxation, Subgradient/Bundle Methods

Algorithms for the solution of linear programs

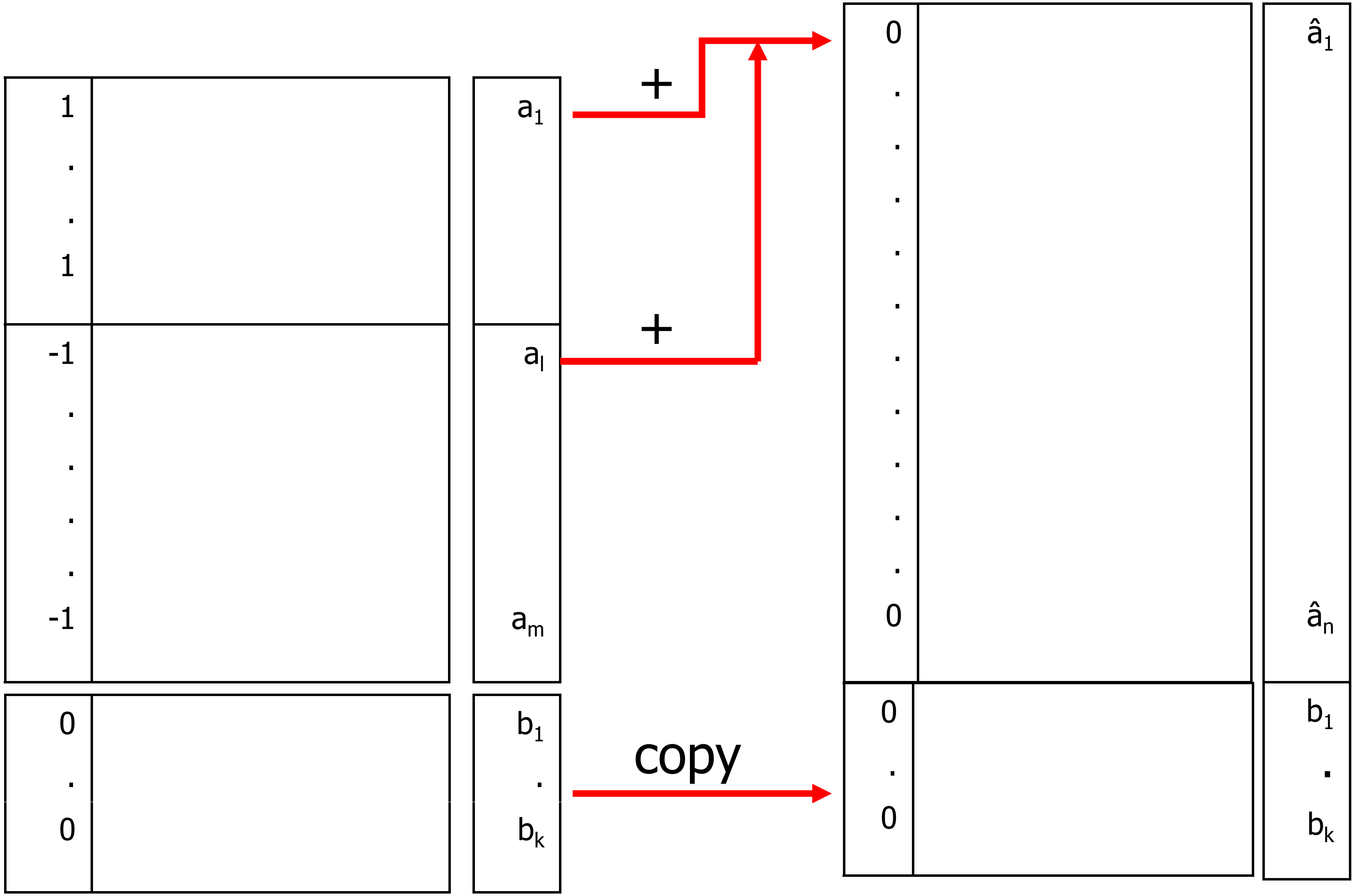
1. Fourier-Motzkin Elimination
2. The Primal Simplex Method
3. The Dual Simplex Method
4. The Ellipsoid Method
5. Interior-Point/Barrier Methods
6. Lagrangian Relaxation, Subgradient/Bundle Methods

Fourier-Motzkin Elimination

- Fourier, 1826/1827
- Motzkin, 1938
- **Method**: successive projection of a polyhedron in n-dimensional space into a vector space of dimension n-1 by elimination of one variable.

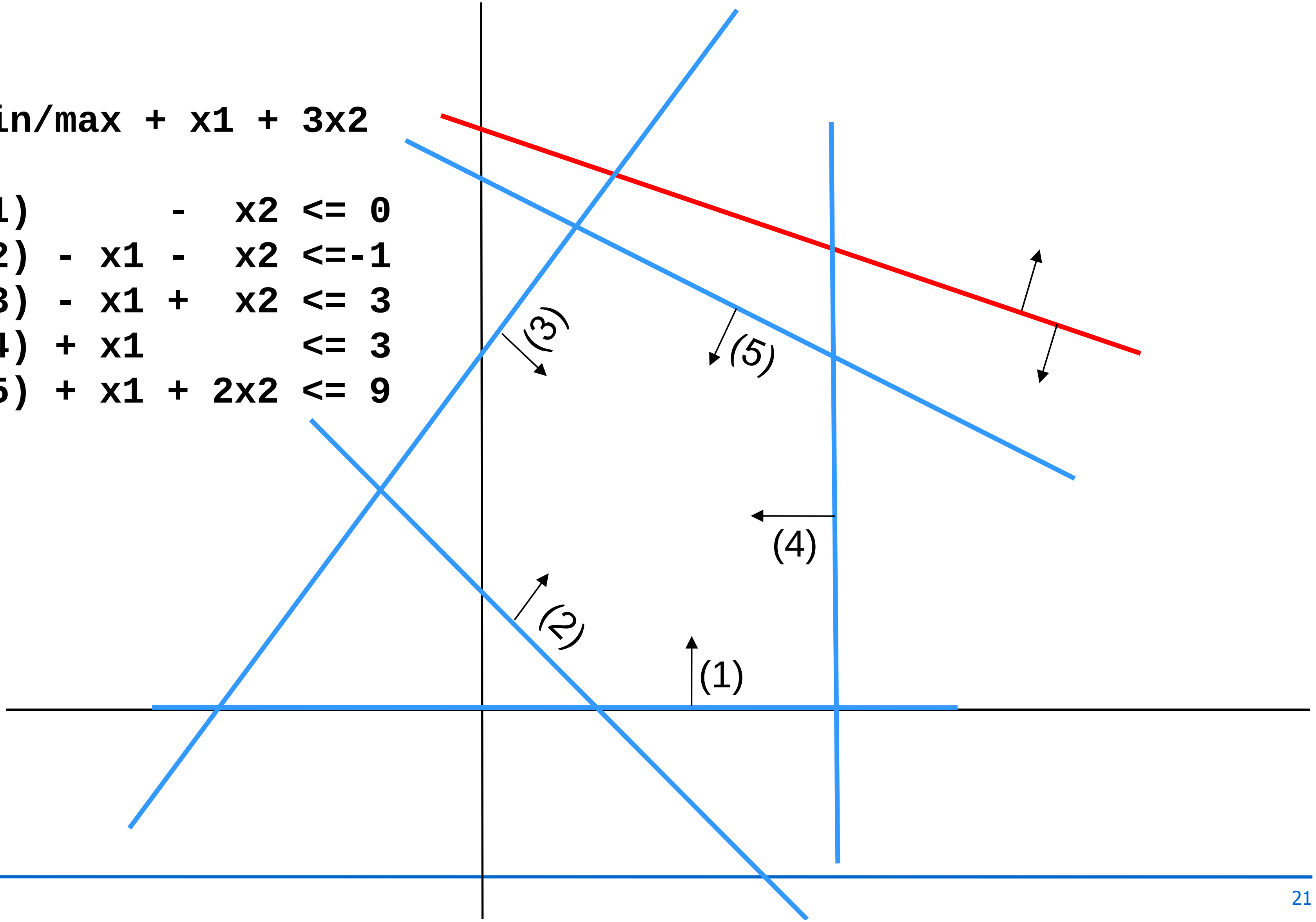


A Fourier-Motzkin step



Fourier-Motzkin Elimination: an example

```
min/max + x1 + 3x2

(1)        -   x2 <= 0
(2) - x1 -   x2 <=-1
(3) - x1 +   x2 <= 3
(4) + x1          <= 3
(5) + x1 + 2x2 <= 9
```



Fourier-Motzkin Elimination: an example, call of PORTA (Polymake)

```
min/max + x1 + 3x2

(1)       -   x2 <= 0
(2) - x1 -   x2 <=-1
(3) - x1 +   x2 <= 3
(4) + x1         <= 3
(5) + x1 + 2x2 <= 9
```

```
DIM = 3

INEQUALITIES_SECTION

(1)       -   x2       <= 0
(2) - x1 -   x2       <=-1
(3) - x1 +   x2       <= 3
(4) + x1               <= 3
(5) + x1 + 2x2        <= 9
(6) + x1 + 3x2 - x3 <= 0
(7) - x1 - 3x2 + x3 <= 0


ELIMINATION_ORDER
1 0 0
```

Fourier-Motzkin Elimination: an example

```
DIM = 3

INEQUALITIES_SECTION

(1)    (1) - x2         <=  0
(2,4)  (2) - x2         <=  2
(2,5)  (3) + x2         <=  8
(2,6)  (4) +2x2 - x3    <= -1
(3,4)  (5) + x2         <=  6
(3,5)  (6) + x2         <=  4
(3,6)  (7) +4x2 - x3    <=  3
(7,4)  (8) -3x2 + x3    <=  3
(7,5)  (9) - x2 + x3    <=  9
(7,6)
```

```
DIM = 3

INEQUALITIES_SECTION

(1)        -  x2        <= 0
(2) - x1 -  x2          <=-1
(3) - x1 +  x2          <= 3
(4) + x1                <= 3
(5) + x1 + 2x2          <= 9
(6) + x1 + 3x2 - x3     <= 0
(7) - x1 - 3x2 + x3     <= 0

ELIMINATION_ORDER
1 0 0
```

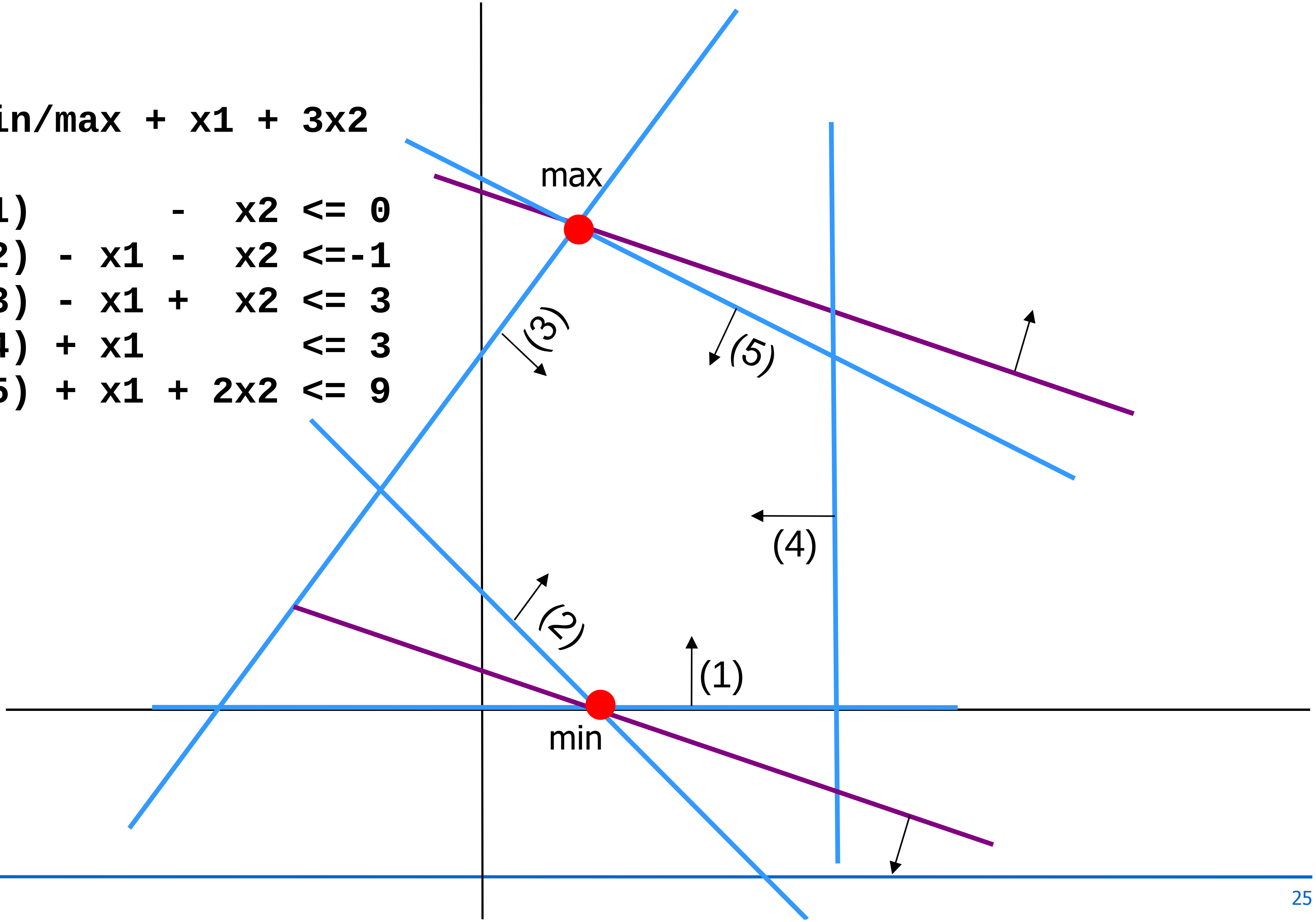
Fourier-Motzkin Elimination: an example

```
DIM = 3

INEQUALITIES_SECTION

(1)   (1) - x2        <=  0
(2,4) (2) - x2        <=  2
(2,5) (3) + x2        <=  8
(2,6) (4) +2x2 - x3 <= -1
(3,4) (5) + x2        <=  6
(3,5) (6) + x2        <=  4
(3,6) (7) +4x2 - x3 <=  3
(7,4) (8) -3x2 + x3 <=  3
(7,5) (9) - x2 + x3 <=  9
(7,6)

ELIMINATION_ORDER
0 1 0
```

```
(1,4) (  1) -x3 <=  -1
(1,7) (  2) -x3 <=   3
(2,4) (  3) -x3 <=   3
(2,7) (  4) -x3 <=  11
(8,3) (  5) +x3 <=  27
(8,4) (  6) -x3 <=   3
(8,5) (  7) +x3 <=  21
(8,6) (  8) +x3 <=  15
(8,7) (  9) +x3 <=  21
(9,3) ( 10) +x3 <=  17
(9,4) ( 11) +x3 <=  17
(9,5) ( 12) +x3 <=  15
(9,6) ( 13) +x3 <=  13
(9,7) ( 14)+3x3 <=  39

min = 1 <= x3 <= 13 = max

x1 = 1            x1 = 1
x2 = 0            x2 = 4
```

Fourier-Motzkin Elimination: an example

```
min/max + x1 + 3x2

(1)        -   x2 <= 0
(2) - x1 -   x2 <=-1
(3) - x1 +   x2 <= 3
(4) + x1        <= 3
(5) + x1 + 2x2 <= 9
```



Fourier-Motzkin Elimination

FME is a wonderful constructive proof method.

Elimination of all variables of a given inequality system directly yields the Farkas Lemma:

$Ax \leq b$ has a solution or

$y^T A = 0, y^T b < 0$ has a solution

but not both.

FME is computationally lousy.

Algorithms for the solution of linear programs

1. Fourier-Motzkin Elimination
2. The Primal Simplex Method
3. The Dual Simplex Method
4. The Ellipsoid Method
5. Interior-Point/Barrier Methods
6. Lagrangian Relaxation, Subgradient/Bundle Methods

The Primal Simplex Method

Dantzig, 1947: primal Simplex Method

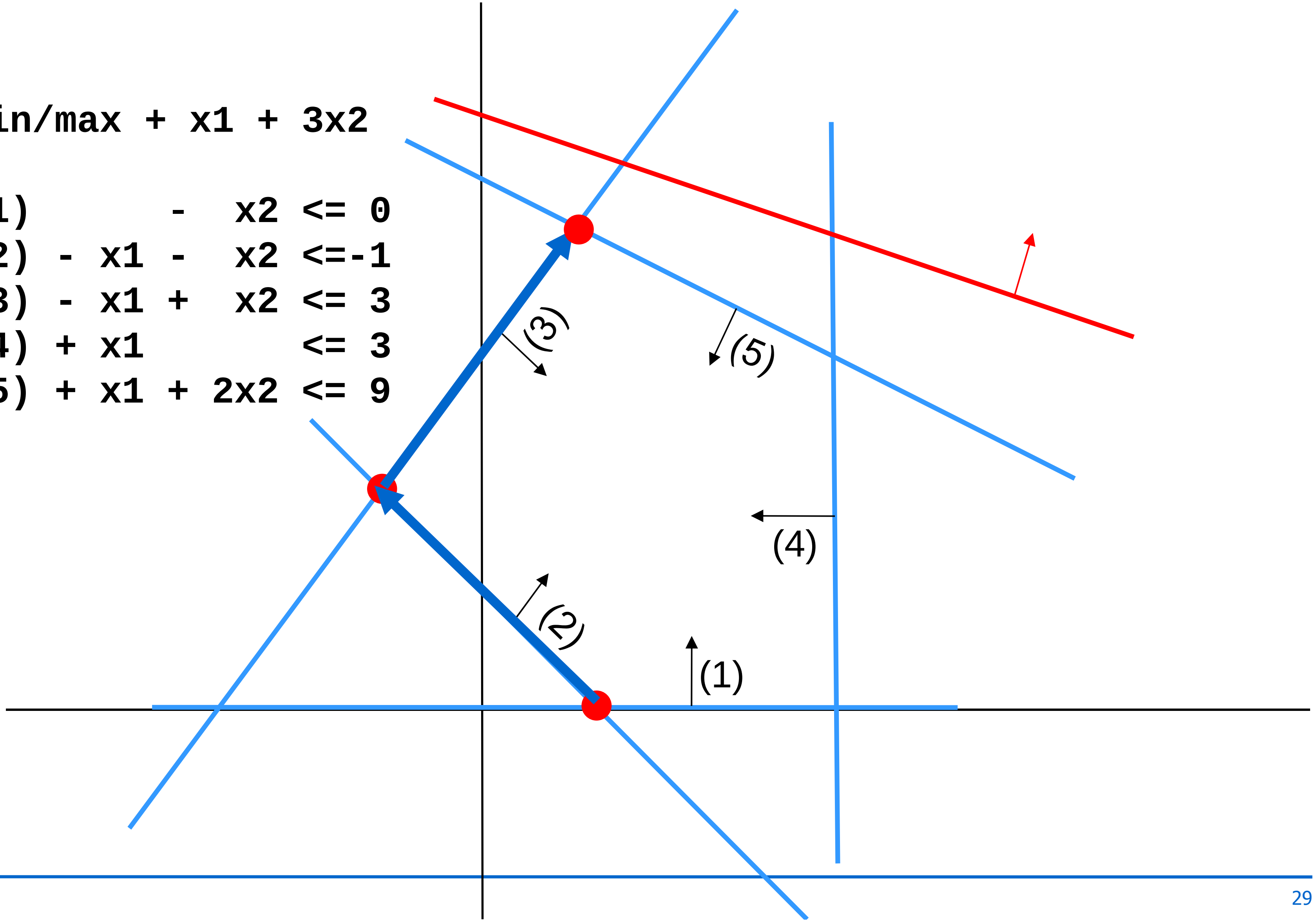
Dantzig, 1953: revised Simplex Method

….

Underlying Idea: Find a vertex of the set of feasible LP solutions (polyhedron) and move to a better neighbouring vertex, if possible (Fourier‘s idea 1826/27).

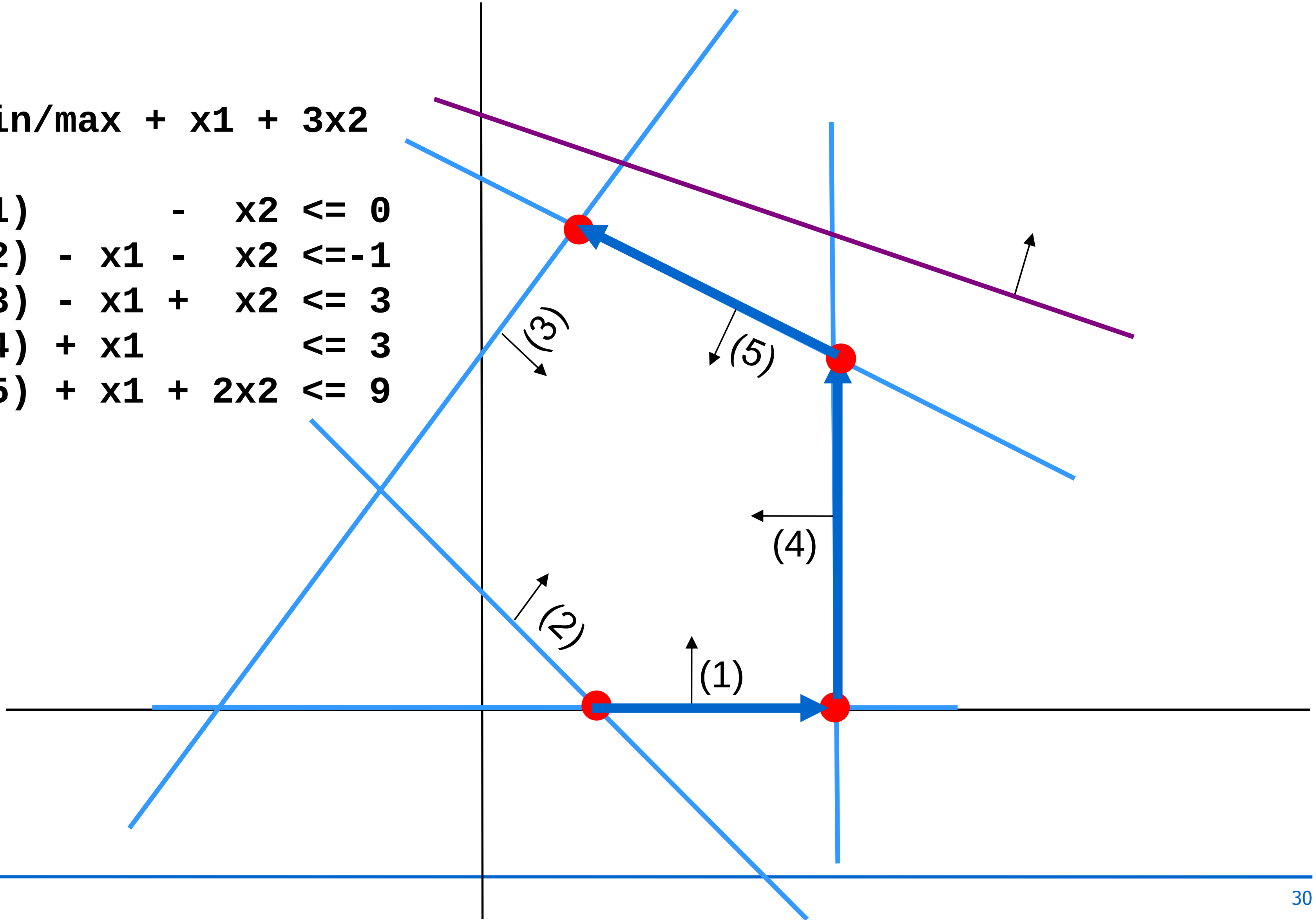
The Simplex Method: an example

```
min/max + x1 + 3x2

(1)        -   x2 <= 0
(2) - x1 -   x2 <=-1
(3) - x1 +   x2 <= 3
(4) + x1        <= 3
(5) + x1 + 2x2 <= 9
```



The Simplex Method: an example

```
min/max + x1 + 3x2

(1)        -   x2 <= 0
(2) - x1 -   x2 <=-1
(3) - x1 +   x2 <= 3
(4) + x1          <= 3
(5) + x1 + 2x2 <= 9
```



Hirsch Conjecture

If P is a polytope of dimension n with m facets then every vertex of P can be reached from any other vertex of P on a path of length at most m-n.

In the example before: m=5, n=2 and m-n=3,

2 or 3 steps are needed, and the conjecture holds (precisely).

At present, not even a polynomial bound on the path length is known.

Disproof:

Santos, Francisco (2011), "A counterexample to the Hirsch conjecture", Annals of Mathematics 176 (1): 383–412, arXiv:1006.2814, doi:10.4007/annals.2012.176.1.7, MR 2925387

Computationally important idea of the Simplex Method

Let a (m,n)-Matrix A with full row rank m, an m-vector b and an n-vector c with m<n be given. For every vertex y of the polyhedron of feasible solutions of the LP,

$$\max c^T x \qquad Ax = b \qquad x \geq 0$$

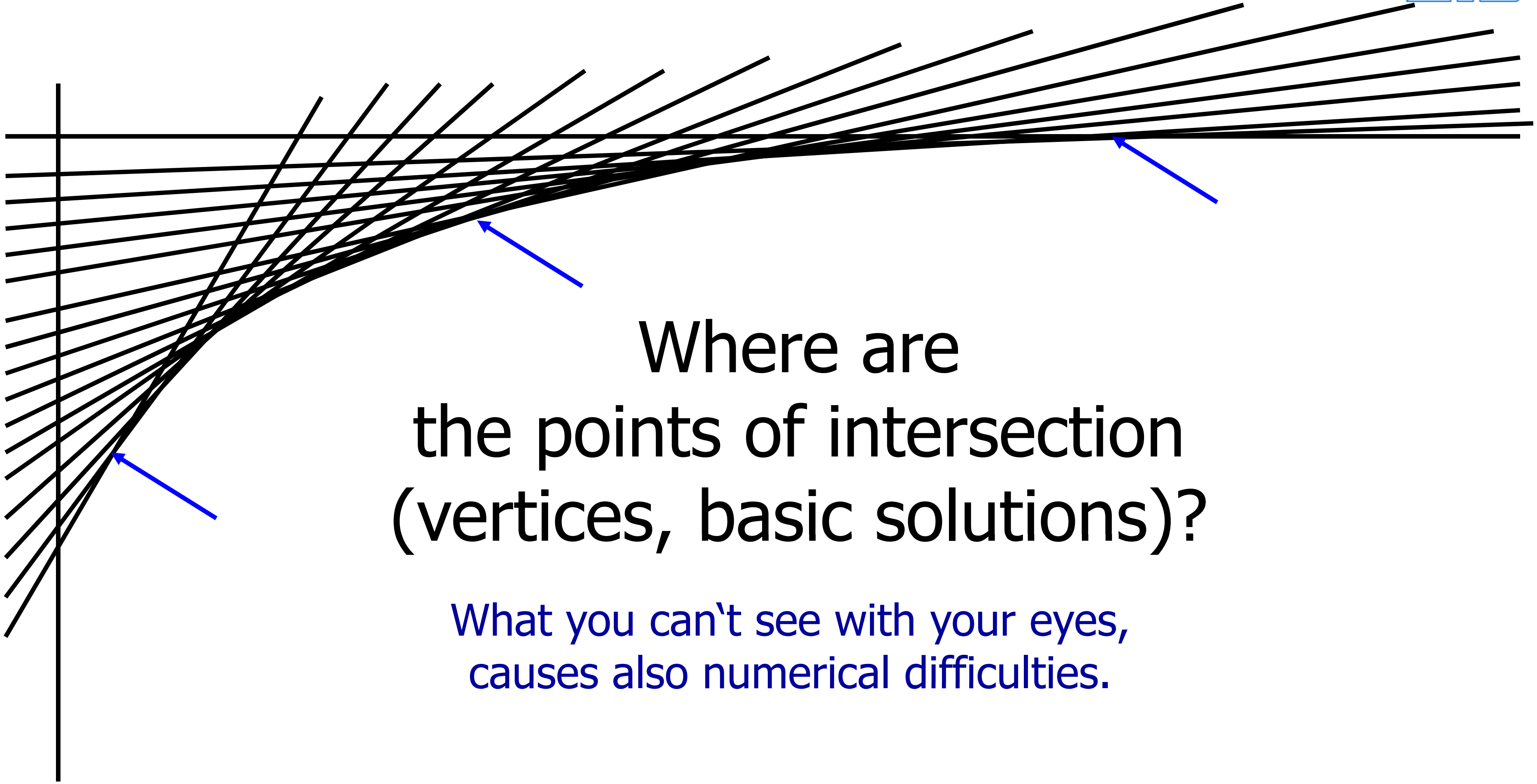
$$A = \begin{array}{|c|c|} \hline \quad B \quad & \qquad\qquad N \qquad\qquad \\ \hline \end{array}$$

there is a non-singular (m,m)-submatrix B (called basis) of A representing the vertex y (basic solution) as follows

$$y_B = B^{-1}b, \quad y_N = 0$$

Many computational consequences:
Update-formulas, reduced cost calculations,
number of non-zeros of a vertex,...

Numerical trouble often has geometric reasons



Where are
the points of intersection
(vertices, basic solutions)?

What you can't see with your eyes,
causes also numerical difficulties.

Algorithms for the solution of linear programs

1. Fourier-Motzkin Elimination
2. The Primal Simplex Method
3. The Dual Simplex Method
4. The Ellipsoid Method
5. Interior-Point/Barrier Methods
6. Lagrangian Relaxation, Subgradient/Bundle Methods

The Dual Simplex Method

Dantzig, 1947: primal Simplex Method

Dantzig, 1953: revised Simplex Method

Lemke, 1954; Beale, 1954: dual Simplex Method

Optimizers‵ dream: Duality theorems

The Duality Theorem of Linear Programming

$$
\begin{array}{lcl}
\max c^T x & = & \min y^T b \\
Ax \le b & & y^T A \ge c^T \\
x \ge 0 & & y \ge 0
\end{array}
$$

Optimizers‵ dream: Duality theorems

Max-Flow Min-Cut Theorem

The value of a maximal (s,t)-flow in a capacitated network is equal to the minimal capacity of an (s,t)-cut.

The Duality Theorem of Linear Programming

$$\begin{array}{lcl} \max c^T x & = & \min y^T b \\ Ax \le b & & y^T A \ge c^T \\ x \ge 0 & & y \ge 0 \end{array}$$

Optimizers‵ dream: Duality theorems for integer programming

The Max-Flow Min-Cut Theorem
does not hold if several source-sink relations are given
(multicommodity flow).

The Duality Theorem of Linear Programming
does not hold if integrality conditions are added

Important technique:

Use polyhedral theory
to obtain “= ”.

$$\max c^T x \quad \leq \quad \min y^T b$$

$$Ax \leq b \qquad \ll \qquad y^T A \geq c^T$$

$$x \geq 0 \qquad\qquad y \geq 0$$

$$x \in \mathbf{Z}^n \qquad\qquad y \in \mathbf{Z}^m$$

Dual Simplex Method

The Dual Simplex Method is the (Primal) Simplex Method applied to the dual of a given linear program.

Surprise in the mid-nineties:

The Dual Simplex Method is faster than the Primal in practice.
One key: Goldfarb's steepest edge pivoting rule!

A wonderful observation for the cutting plane methods of integer programming!

Ask Bob Bixby for a detailed explanation!

Algorithms for the solution of linear programs

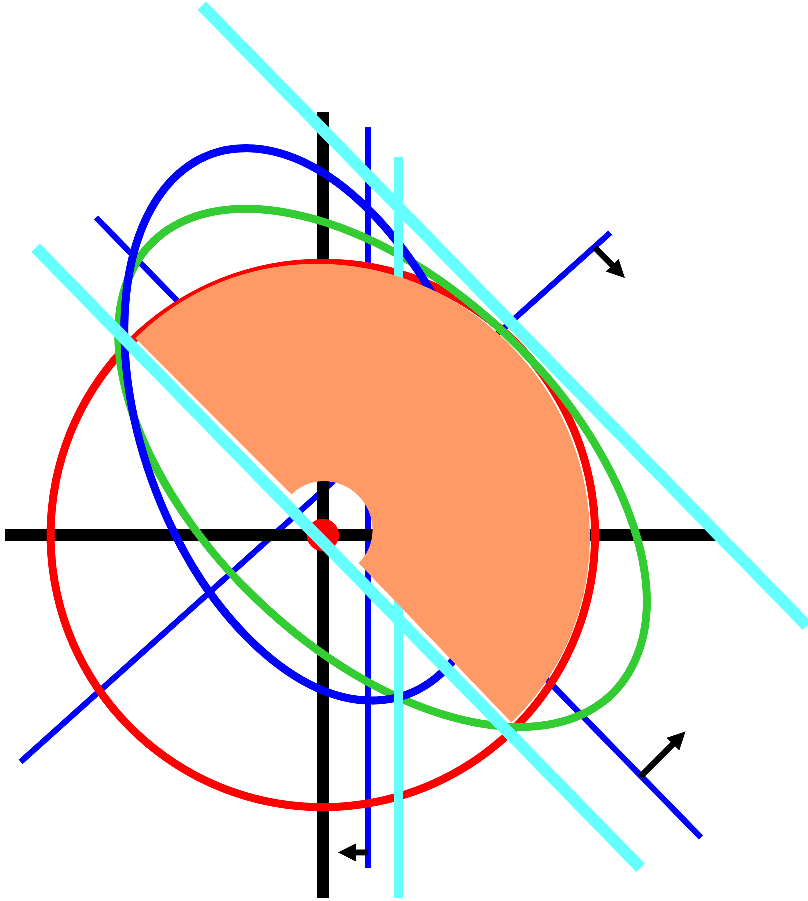
1. Fourier-Motzkin Elimination
2. The Primal Simplex Method
3. The Dual Simplex Method
4. The Ellipsoid Method
5. Interior-Point/Barrier Methods
6. Lagrangian Relaxation, Subgradient/Bundle Methods

The Ellipsoid Method

- Shor, 1970 - 1979
- Yudin & Nemirovskii, 1976
- Khachiyan, 1979
- M. Grötschel, L. Lovász, A. Schrijver,
  Geometric Algorithms and Combinatorial Optimization
  Algorithms and Combinatorics 2, Springer, 1988

  This book can be downloaded from my homepage!
- http://www.zib.de/groetschel/pubnew/paper/groetschellovaszschrijver1988.pdf

The Ellipsoid Method: an example



$$k := 0,$$
$$N := 2n((2n+1)\langle C\rangle + n\langle d\rangle - n^3)$$
$$A_0 := R^2 I \quad \text{with } R := \sqrt{n}\, 2^{\langle C,d\rangle - n^2}$$

$$\boxed{P := \{x \mid Cx \leq d\}}$$

Initialization

$$a_0 := 0$$

If $k = N$, *STOP!* (Declare P empty.)
If $a_k \in P$, *STOP!* (A feasible solution is found.)
If $a_k \notin P$, *then choose an inequality, say* $c^T x \leq \gamma$, *of the system* $Cx \leq d$ *that is violated by* a_k.

Stopping criterion
Feasibility check
Cutting plane choice

$$b := \frac{1}{\sqrt{c^T A_k c}} A_k c$$

$$a_{k+1} := a_k - \frac{1}{n+1} b$$

Update

$$A_{k+1} := \frac{n^2}{n^2-1}\left(A_k - \frac{2}{n+1} b b^T\right)$$

The Ellipsoid Method

Доклады Академии наук СССР
1979. Том 244, № 5

УДК 519.95

МАТЕМАТИКА

Л. Г. ХАЧИЯН

ПОЛИНОМИАЛЬНЫЙ АЛГОРИТМ В ЛИНЕЙНОМ ПРОГРАММИРОВАНИИ

(Представлено академиком А. А. Дородницыным 4 X 1978)

Рассмотрим систему из $m \geqslant 2$ линейных неравенств относительно $n \geqslant 2$ вещественных переменных $x_1, \ldots, x_j, \ldots, x_n$

$$a_{i1}x_1 + \ldots + a_{in}x_n \leqslant b_i, \quad i=1, 2, \ldots, m, \tag{1}$$

с целыми коэффициентами a_{ij}, b_i. Пусть

$$L = \left[\sum_{i,j=1}^{m,n} \log_2(|a_{ij}|+1) + \sum_{i=1}^{m} \log_2(|b_i|+1) + \log_2 nm \right] + 1 \tag{2}$$

есть длина входа системы, т. е. число символов 0 и 1, необходимых для записи (1) в двоичной системе счисления.

May 1980, New York Times, Nov. 7, 1979

A Soviet Discovery Rocks World of Mathematics
By MALCOLM W. BROWNE
New York Times (1857-Current file); Nov 7, 1979; ProQuest Historical Newspapers The New York Times (1851 - 2003)
pg. A1

A Soviet Discovery Rocks World of Mathematics

By MALCOLM W. BROWNE

A surprise discovery by an obscure Soviet mathematician has rocked the world of mathematics and computer analysis, and experts have begun exploring its practical applications.

Mathematicians describe the discovery by L.G. Khachian as a method by which computers can find guaranteed solutions to a class of very difficult problems that have hitherto been tackled on a kind of hit-or-miss basis.

Apart from its profound theoretical interest, the discovery may be applicable in weather prediction, complicated industrial processes, petroleum refining, the scheduling of workers at large factories, secret codes and many other things.

"I have been deluged with calls from virtually every department of government for an interpretation of the significance of this," a leading expert on computer methods, Dr. George B. Dantzig of Stanford University, said in an interview.

The solution of mathematical problems by computer must be broken down into a series of steps. One class of problem sometimes involves so many steps that it could take billions of years to compute.

The Russian discovery offers a way by which the number of steps in a solution can be dramatically reduced. It also offers the mathematician a way of learning quickly whether a problem has a solution or not, without having to complete the entire immense computation that may be required.

According to the American journal Sci-

Continued on Page A20, Column 3

May 1980
National Association of Science Writers

NEWSLETTER

NATIONAL ASSOCIATION OF SCIENCE WRITERS

Volume 28, Number 2 May 1980

Science Writers Rock World of Mathematics: Tales of the Traveling Salesman Problem

by Jonathan Weiner

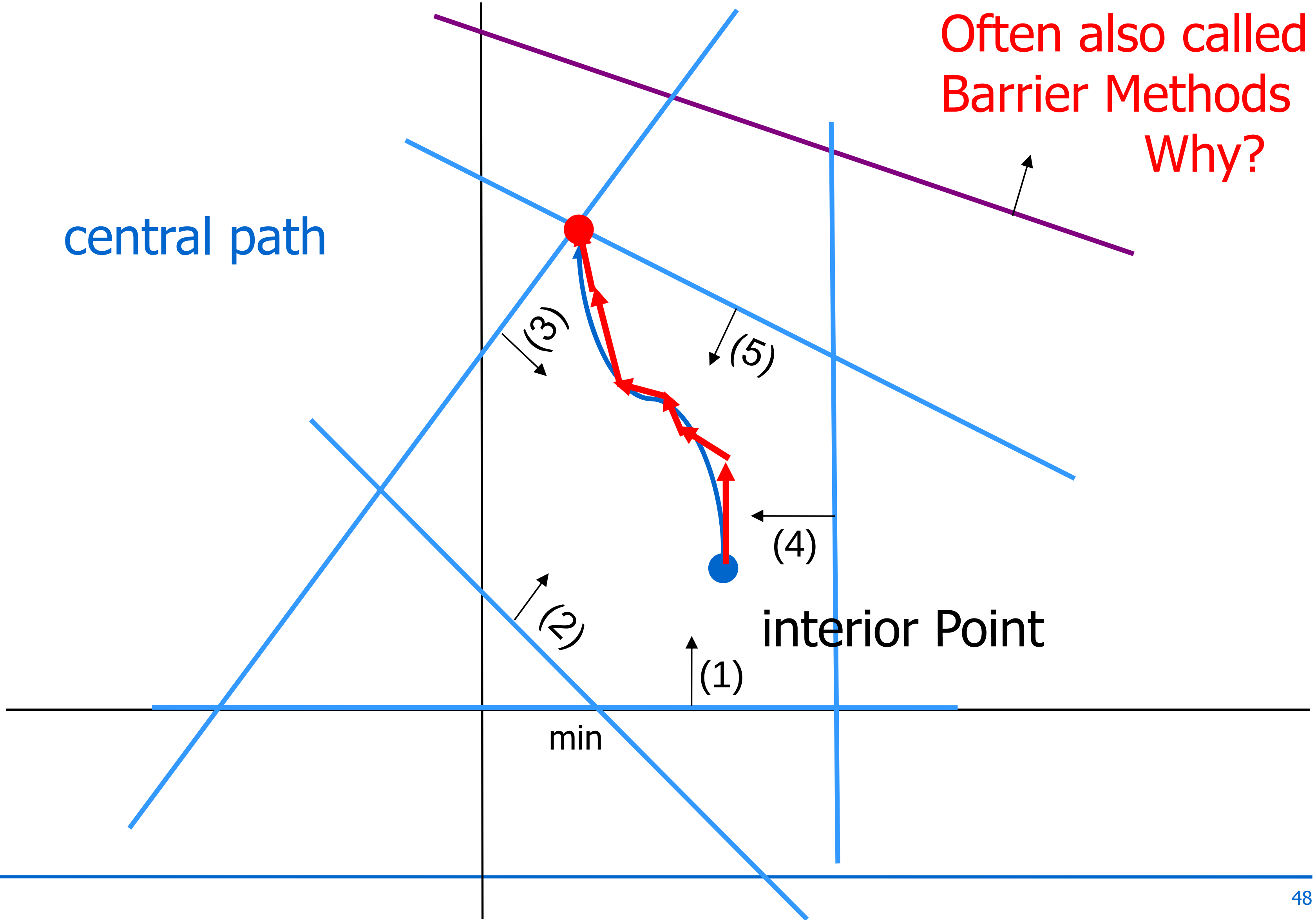
Echoes of Sputnik. An obscure young Russian mathematician solves a key problem in linear programming, and American defense experts wring their hands worrying about its applications to secret codes, weather forecasting, and Kremlin-only-knows what else.

It was a pretty good story, as mathematics news goes, and it wound up on page one of *The New York Times* last November 7. It was run by *The Times* news service, and it was picked up far and wide as a nifty science novelty item. It had all the elements of a spy novel: the cold war, a valuable scientific formula, and sexual innuendo in the form of a traveling salesman. Who could ignore it?

Algorithms for the solution of linear programs

1. Fourier-Motzkin Elimination
2. The Primal Simplex Method
3. The Dual Simplex Method
4. The Ellipsoid Method
5. Interior-Point/Barrier Methods
6. Lagrangian Relaxation, Subgradient/Bundle Methods

Interior-Point Methods: an example



The Karmarkar Algorithm

(13.25) Der Karmarkar-Algorithmus.

Input: $A \in \mathbb{Q}^{(m,n)}$ *und* $c \in \mathbb{Q}^n$. *Zusätzlich wird vorausgesetzt, dass* $\frac{1}{n}A\mathbb{1} = 0$ *und* $c^T\mathbb{1} > 0$ *gilt.*

Output: *Ein Vektor* x *mit* $Ax = 0$, $\mathbb{1}^T x = 1$, $x \geq 0$ *und* $c^T x \leq 0$ *oder die Feststellung, dass kein derartiger Vektor existiert.*

(1) **Initialisierung.** *Setze*

$$\begin{aligned} x^0 &:= \tfrac{1}{n}\mathbb{1} \\ k &:= 0 \\ N &:= 3n(\langle A\rangle + 2\langle c\rangle - n) \end{aligned}$$

(2) **Abbruchkriterium.**

(2.a) *Gilt* $k = N$, *dann hat* $Ax = 0$, $\mathbb{1}^T x = 1$, $x \geq 0$, $c^T x \leq 0$ *keine Lösung, STOP!*

(2.b) *Gilt* $c^T x^k \leq 2^{-\langle A\rangle - \langle c\rangle}$, *dann ist eine Lösung gefunden. Falls* $c^T x^k \leq 0$, *dann ist* x^k *eine Lösung, andernfalls kann wie bei der Ellipsoidmethode (Satz (12.34)) aus* x^k *ein Vektor* $\overline{x}$ *konstruiert werden mit* $c^T\overline{x} \leq 0$ $A\overline{x} = 0$, $\mathbb{1}^T\overline{x} = 1$, $\overline{x} \geq 0$, *STOP!*

Update.

(3) *(3.a)* $D := \operatorname{diag}(x^k)$

(3.b) $\overline{\overline{c}} := (I - DA^T(AD^2A^T)^{-1}AD - \frac{1}{n}\mathbb{1}\mathbb{1}^T)Dc$

(3.c) $y^{k+1} := \frac{1}{n}\mathbb{1} - \frac{1}{2}\frac{1}{\sqrt{n(n-1)}}\frac{1}{\|\overline{\overline{c}}\|}\overline{\overline{c}}$

(3.d) $x^{k+1} := \frac{1}{\mathbb{1}^T D y^{k+1}} D y^{k+1}$

(3.e) $k := k + 1$

Gehe zu (2).

Breakthrough in Problem Solving

By JAMES GLEICK

A 28-year-old mathematician at A.T.&T. Bell Laboratories has made a startling theoretical breakthrough in the solving of systems of equations that often grow too vast and complex for the most powerful computers.

The discovery, which is to be formally published next month, is already circulating rapidly through the mathematical world. It has also set off a deluge of inquiries from brokerage houses, oil companies and airlines, industries with millions of dollars at stake in problems known as linear programming.

Faster Solutions Seen

These problems are fiendishly complicated systems, often with thousands of variables. They arise in a variety of commercial and government applications, ranging from allocating time on a communications satellite to routing millions of telephone calls over long distances, or whenever a limited, expensive resource must be spread most efficiently among competing users. And investment companies use them in creating portfolios with the best mix of stocks and bonds.

The Bell Labs mathematician, Dr. Narendra Karmarkar, has devised a radically new procedure that may speed the routine handling of such problems by businesses and Government agencies and also make it possible to tackle problems that are now far out of reach.

"This is a path-breaking result," said Dr. Ronald L. Graham, director of mathematical sciences for Bell Labs in Murray Hill, N.J. "Science has its moments of great progress, and this may well be one of them."

Because problems in linear programming can have billions or more possible answers, even high-speed computers cannot check every one. So computers must use a special procedure, an algorithm, to examine as few answers as possible before finding the best one — typically the one that minimizes cost or maximizes efficiency.

A procedure devised in 1947, the simplex method, is now used for such problems,

Continued on Page A19, Column 1

THE NEW YORK TIMES, November 19, 1984

Karmarkar at Bell Labs: an equation to find a new way through the maze

Folding the Perfect Corner

A young Bell scientist makes a major math breakthrough

Every day 1,200 American Airlines jets crisscross the U.S., Mexico, Canada and the Caribbean, stopping in 110 cities and bearing over 80,000 passengers. More than 4,000 pilots, copilots, flight personnel, maintenance workers and baggage carriers are shuffled among the flights; a total of 3.6 million gal. of high-octane fuel is burned. Nuts, bolts, altimeters, landing gears and the like must be checked at each destination. And while performing these scheduling gymnastics, the company must keep a close eye on costs, projected revenue and profits.

Like American Airlines, thousands of companies must routinely untangle the myriad variables that complicate the efficient distribution of their resources. Solving such monstrous problems requires the use of an abstruse branch of mathematics known as linear programming. It is the kind of math that has frustrated theoreticians for years, and even the fastest and most powerful computers have had great difficulty juggling the bits and pieces of data. Now Narendra Karmarkar, a 28-year-old Indian-born mathematician at Bell Laboratories in Murray Hill, N.J., after only a years' work has cracked the puzzle of linear programming by devising a new algorithm, a step-by-step mathematical formula. He has translated the procedure into a program that should allow computers to track a greater combination of tasks than ever before and in a fraction of the time.

Unlike most advances in theoretical mathematics, Karmarkar's work will have an immediate and major impact on the real world. "Breakthrough is one of the most abused words in science," says Ronald Graham, director of mathematical sciences at Bell Labs. "But this is one situation where it is truly appropriate."

Before the Karmarkar method, linear equations could be solved only in a cumbersome fashion, ironically known as the simplex method, devised by Mathematician George Dantzig in 1947. Problems are conceived of as giant geodesic domes with thousands of sides. Each corner of a facet on the dome

TIME MAGAZINE, December 3, 1984

Milestones for Interior Point Methods (IPMs)

- 1984 Projective IPM: Karmarkar – efficient in practice!?
- 1989 $O(n^3L)$ for IPMs: Renegar – best complexity
- 1989 Primal–Dual IPMs: Kojima ... – dominant since then
- 1989 Self-Concordant Barrier: Nesterov–Nemirovskii – extensions to smooth convex optimozation
- 1992 Semi-Definite Optimization (SDO) and Second Order Conic Optimization (SOCO): Alizadeh, Nesterov–Nemirovskii –new applications, approximations, software
- 1998 Robust LO: Ben Tal–Nemirovskii

Complexity of Self-Regular IPMs

Method	Large update	Small update	
θ	$1 - 1/100$	$1/\sqrt{n}$	
Iter. bound	$\mathcal{O}(n \log \frac{n}{\epsilon})$	$\mathcal{O}(\sqrt{n} \log \frac{n}{\epsilon})$	
Performance	Efficient	Very poor	
SR-Method	SR-Large	SR-Small	SR-Large $q = \log n$
θ	$1 - 1/100$	$1/\sqrt{n}$	constant
Iter. bound	$\mathcal{O}(q\, n^{\frac{q+1}{2q}} \log \frac{n}{\epsilon})$	$\mathcal{O}(\sqrt{n} \log \frac{n}{\epsilon})$	$\mathcal{O}(\sqrt{n} \log n \log \frac{n}{\epsilon})$
Performance	Efficient	Very poor	Efficient

"Almost" constant (< 100) number of iterations in practice!

Algorithms for the solution of linear programs

1. Fourier-Motzkin Elimination
2. The Primal Simplex Method
3. The Dual Simplex Method
4. The Ellipsoid Method
5. Interior-Point/Barrier Methods
6. Lagrangian Relaxation, Subgradient/Bundle Methods

Lagrangean Relaxation & Non-differentiable Optimization

Approach for very large scale and structured LPs

Methods:

- subgradient
- bundle
- bundle trust region

  or any other nondifferentiable NLP method that looks promissing

Lagrangian Relaxation

Turning an LP into a nonlinear nondifferentiable optimization problem

$$\min c^T x$$
$$Ax = b$$
$$\left.\boxed{\begin{array}{c} Dx \le d \\ x \ge 0 \end{array}}\right. =: Q$$

$$\max f(\lambda)$$
$$f(\lambda) := \min_{x \in Q} c^T x + \lambda^T (Ax - b)$$

(14.25) Satz. *Sei Q nicht leer und endlich und $f(\lambda) := \min_{x \in Q}(c^T x + \lambda^T (Ax - b))$, so gilt folgendes: Setzen wir für $\lambda_0 \in \mathbb{R}^m$, $L_0 := \{x_0 \in \mathbb{R}^m \mid f(\lambda_0) = c^T x_0 + \lambda_0^T (Ax_0 - b)\}$, so ist*

$$\partial f(\lambda_0) = \operatorname{conv}\{(Ax_0 - b) \mid x_0 \in L_0\}\,.$$

Algorithms for nonlinear nondifferential programming

$$x_{i+1} = x_i + s_i d_i$$

$$d_i = \text{subgradient (instead of gradient)}$$

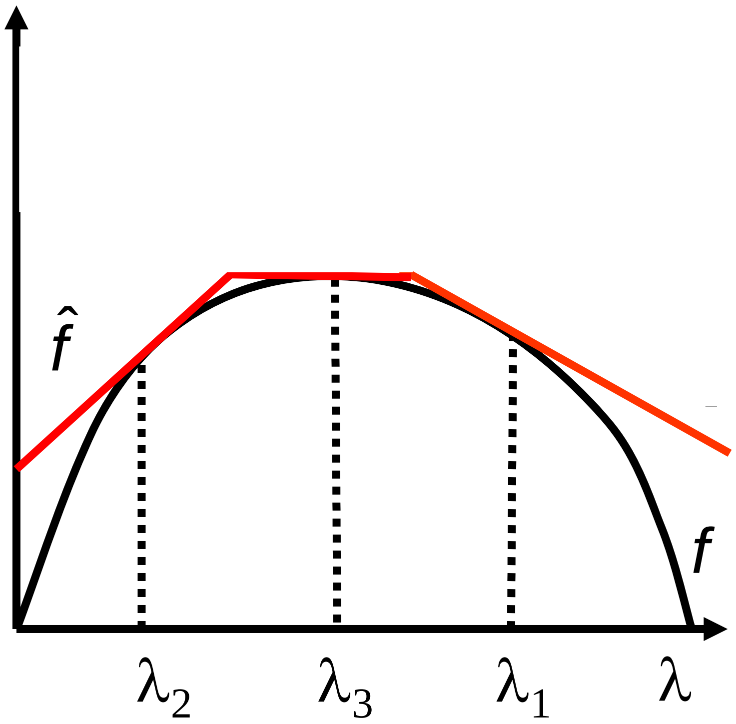
$$\qquad \text{or element of } \varepsilon\text{-subdifferential (bundle)}$$

$$s_i = \text{steplength}$$

Bundle Method
(Kiwiel [1990], Helmberg [2000])

Max $\quad f(\lambda) := \min_{x \in X} c^{\mathrm{T}} x + \lambda^{\mathrm{T}}(b - Ax)$

X polyhedral (piecewise linear)



$$\bar{f}_\mu(\lambda) = c^{\mathrm{T}} x_\mu + \lambda^{\mathrm{T}}(b - Ax_\mu)$$

$$\hat{f}_k(\lambda) := \min_{\mu \in J_k} \bar{f}_\mu(\lambda)$$

$$\lambda_{k+1} = \operatorname*{argmax}_{\lambda} \hat{f}_k(\lambda) - \frac{u_k}{2} \left\| \lambda - \hat{\lambda}_k \right\|^2$$

Quadratic Subproblem

$$(1) \quad \max \hat{f}_k(\lambda) - \frac{u_k}{2}\left\|\lambda - \hat{\lambda}_k\right\|^2$$

$$\Leftrightarrow \quad (2) \quad \begin{aligned} \max \quad & v - \frac{u_k}{2}\left\|\lambda - \hat{\lambda}^k\right\|^2 \\ \text{s.t.} \quad & v \le \bar{f}_\mu(\lambda), \text{ for all } \mu \in J_k \end{aligned}$$

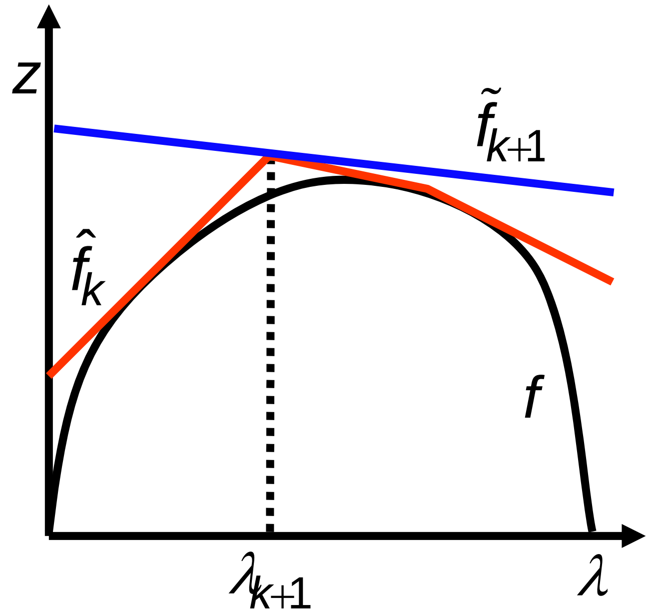
$$\Leftrightarrow \quad (3) \quad \begin{aligned} \max \quad & \sum_{\mu \in J_k} \alpha_\mu \bar{f}_\mu(\hat{\lambda}) - \frac{1}{2u_k}\left\|\sum_{\mu \in J_k} \alpha_\mu (b - Ax_\mu)\right\|^2 \\ \text{s.t.} \quad & \sum_{\mu \in J_k} \alpha_\mu = 1 \\ & 0 \le \alpha_\mu \le 1, \qquad \text{for all } \mu \in J_k \end{aligned}$$

Primal Approximation

$$\lambda_{k+1} = \hat{\lambda}_k + \frac{1}{u} \sum_{\mu \in J_k} \alpha_\mu (b - Ax_\mu)$$

$$\tilde{x}_{k+1} = \sum_{\mu \in J_k} \alpha_\mu x_\mu$$

$$\tilde{f}_k(\lambda) = c^T \tilde{x}_k + \lambda(b - A\tilde{x}_k)$$



Theorem

$$\left\| b - A\tilde{x}_k \right\| \to 0 \ (k \to \infty)$$

$\Rightarrow (\tilde{x}_k)_{k \in N}$ converges to a point $\bar{x} \in \{x : Ax = b, x \in X\}$

Where Bundle Wins

RALF BORNDÖRFER   ANDREAS LÖBEL   STEFFEN WEIDER

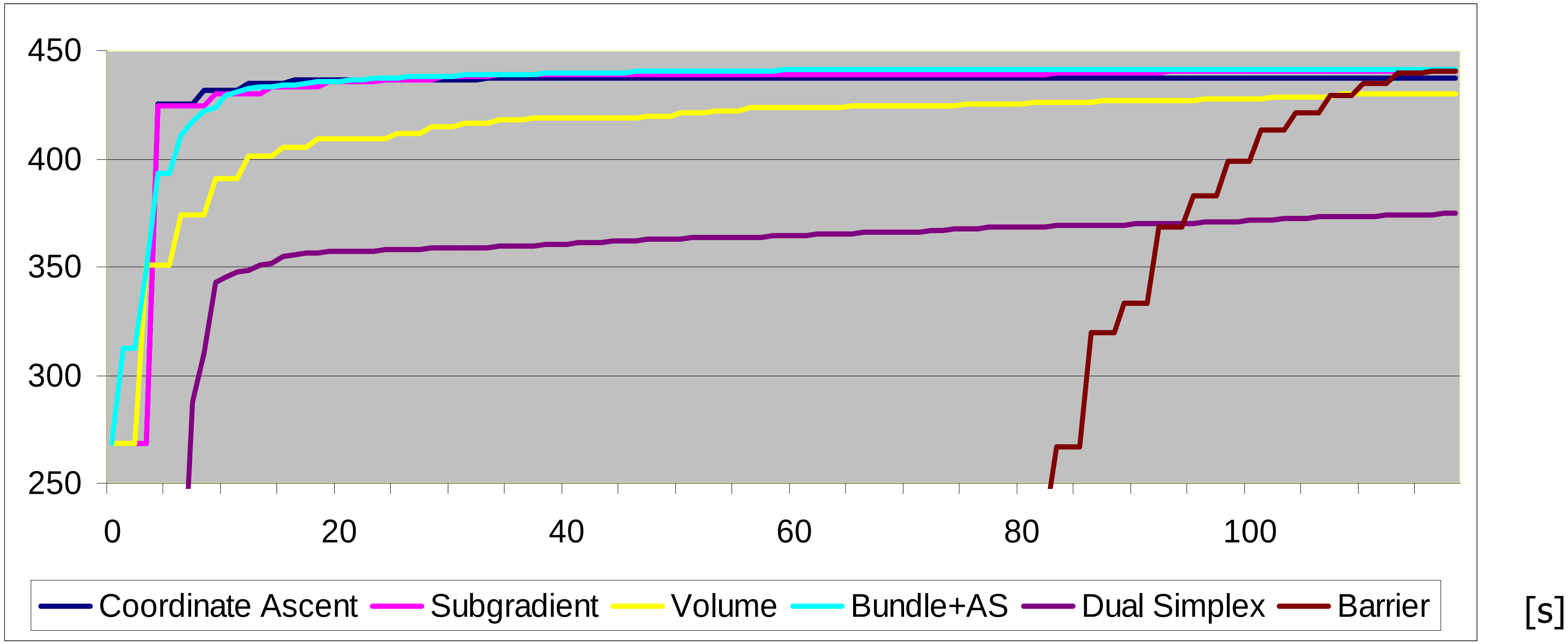
A Bundle Method for Integrated Multi-Depot Vehicle and Duty Scheduling in Public Transit

Computational Results for a (Duty Scheduling) Set Partitioning Model

Duty Scheduling Problem Ivu41:

- 870 500 col
- 3 570 rows
- 10.5 non-zeroes per col

Coordinate Ascent:	Fast, low quality
Subgradient:	(Theoretical) Convergence
Volume:	Primal approximation
Bundle+AS:	Conv. + primal approx.
Dual Simplex:	Primal+dual optimal
Barrier:	Primal+dual optimal



Algorithms for the solution of linear programs

1. Fourier-Motzkin Elimination
2. The Primal Simplex Method
3. The Dual Simplex Method
4. The Ellipsoid Method
5. Interior-Point/Barrier Methods
6. Lagrangian Relaxation, Subgradient/Bundle Methods

Algorithms for the solution of linear programs

1. Fourier-Motzkin Elimination
2. The Primal Simplex Method
3. The Dual Simplex Method
4. The Ellipsoid Method
5. Interior-Point/Barrier Methods
6. Lagrangian Relaxation, Subgradient/Bundle Methods

Conclusions

ZIB Instances

	Variables	Constraints	Non-zeros	Description
1	12,471,400	5,887,041	49,877,768	Group Channel Routing on a 3D Grid Graph (Chip-Bus-Routing)
2	37,709,944	9,049,868	146,280,582	Group Channel Routing on a 3D Grid Graph (different model, infeasible)
3	29,128,799	19,731,970	104,422,573	Steiner-Tree-Packing on a 3D Grid Graph
4	37,423	7,433,543	69,004,977	Integrated WLAN Transmitter Selection and Channel Assignment
5	9,253,265	9,808	349,424,637	Duty Scheduling with base constraints

LP/MIP survey

Robert E. Bixby, Solving Real-World Linear Programs:
A Decade and More of Progress.
Operations Research 50 (2002)3-15.

Newest results at CO@W next week

Which LP solvers are used in practice?

Preview summary

- Fourier-Motzkin: hopeless
- Ellipsoid Method: total failure
- primal Simplex Method: good
- dual Simplex Method: better than primal
- Barrier Method: for large LPs frequently best
- Subgradient Methods: only useful for extremely large scale

- For LP relaxations of IPs: dual Simplex Method

- Who would have predicted that from theoretical insights?

Contents

Mixed Integer Program (MIP)

Characteristics

Objective function
- linear function

Feasible region
- described by linear constraints

Variable domains
- real or integer values

$$\min c^T x$$

$$\text{s.t. } Ax \le b$$

$$(x_I, x_C) \in \mathbb{Z}^I \times \mathbb{R}^{\mathrm{I}}$$

$$c \in \mathbb{R}^n,\; b \in \mathbb{R}^n,\; A \in \mathbb{R}^{m \times n}$$

in reality all numbers are rational

Mixed Integer Program (MIP)

Characteristics

Objective function
- linear function

Feasible region
- described by linear constraints

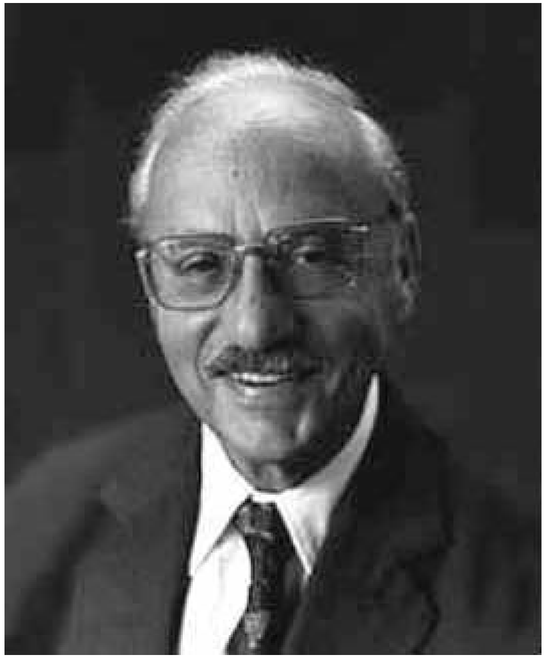
Variable domains
- real or integer values

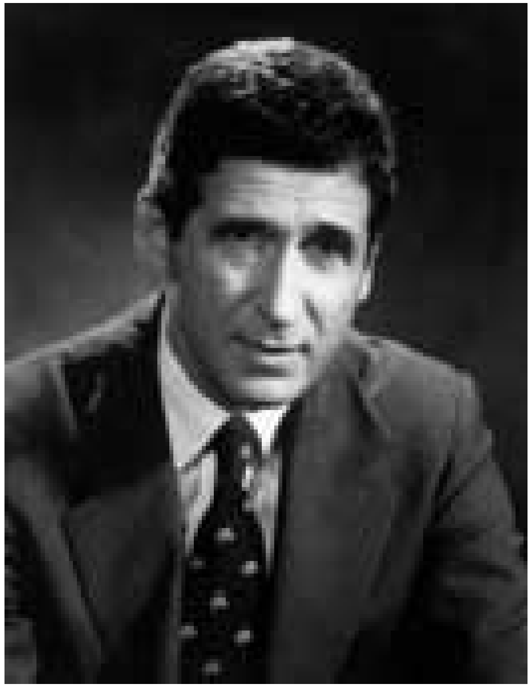
$$\min c^T x$$
$$Ax = a$$
$$Bx \leq b$$
$$x \geq 0$$
$$\text{some } x_j \in \mathbf{Z}$$
$$\text{some } x_j \in \{0,1\}$$

(IP, MIP LMIP, MILP, 0/1-LP, ...)

in reality all numbers are rational

George Dantzig and Ralph Gomory





„founding fathers"

~1950
linear programming

~1960
integer programming

George Dantzig and Ralph Gomory

ISMP Atlanta 2000

Dantzig and Bixby

George Dantzig and
Bob Bixby

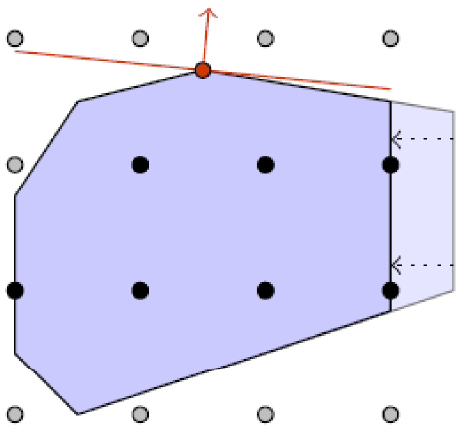
at the International Symposium
on Mathematical Programming,
Atlanta, August 2000

MIP-Solving technologies

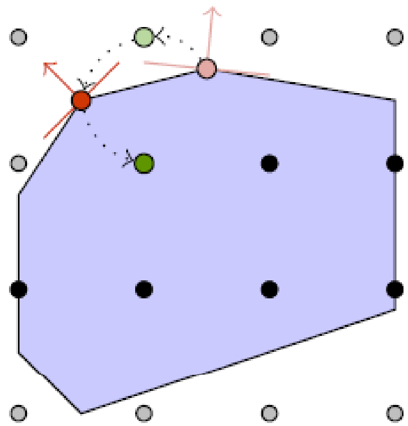
1. Branch and Bound
2. Cutting Planes
3. Column Generation
4. Primal and Dual Heuristics
5. Constraint Programming Ideas

MIP Solver Techniques

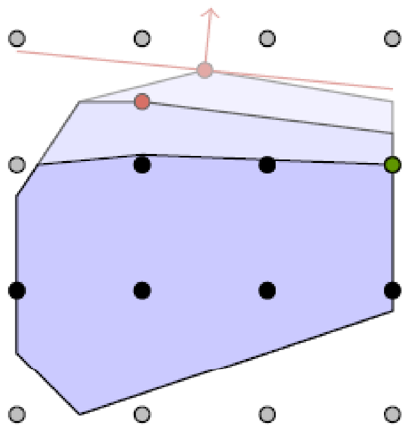
Presolving



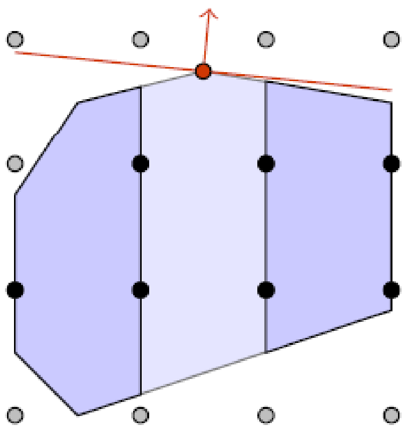
Primal Heuristics



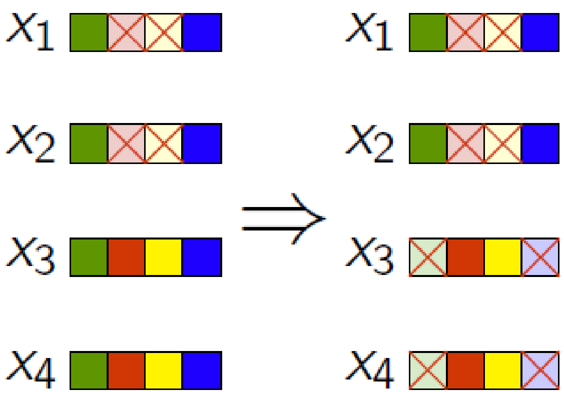
Cutting Planes



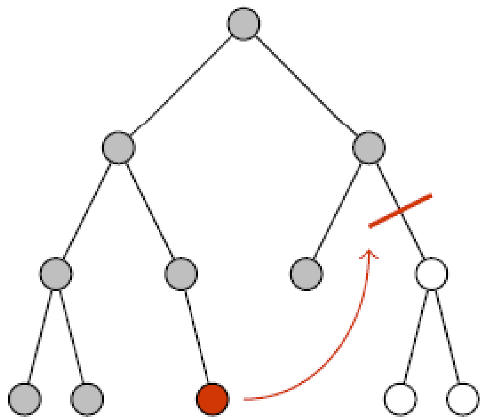
Branch & Bound



Domain Propagation



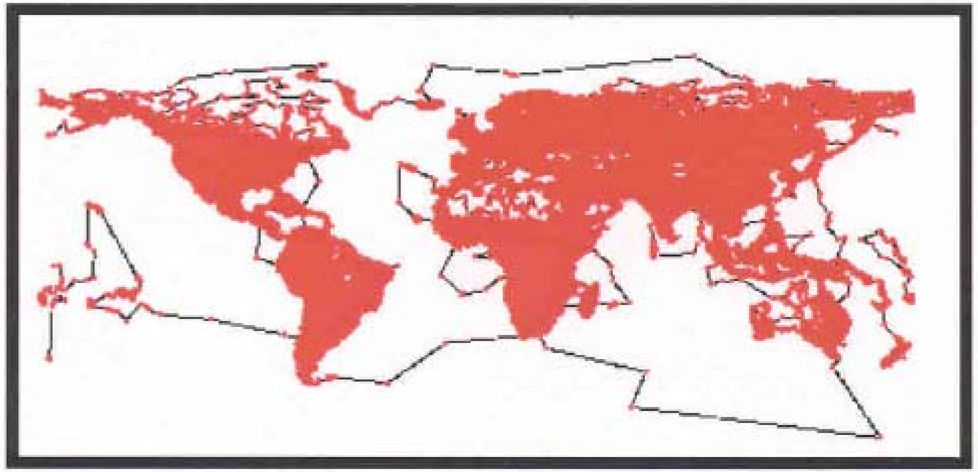
Conflict Analysis



The importance of LP in IP solving
(slide from Bill Cook)

Best current tour length 7,515,772,212 was
found on May 24, 2013, by Keld Helsgaun

1,904,711–City World TSP, 2001



K	Optimality Gap
0	0.235%
8	0.190%
12	0.135%
14	0.111%
16	0.103%

Solution of LP Problems takes over 99% of CPU time

# of variables = 1,813,961,044,405 = 1,8 trillion

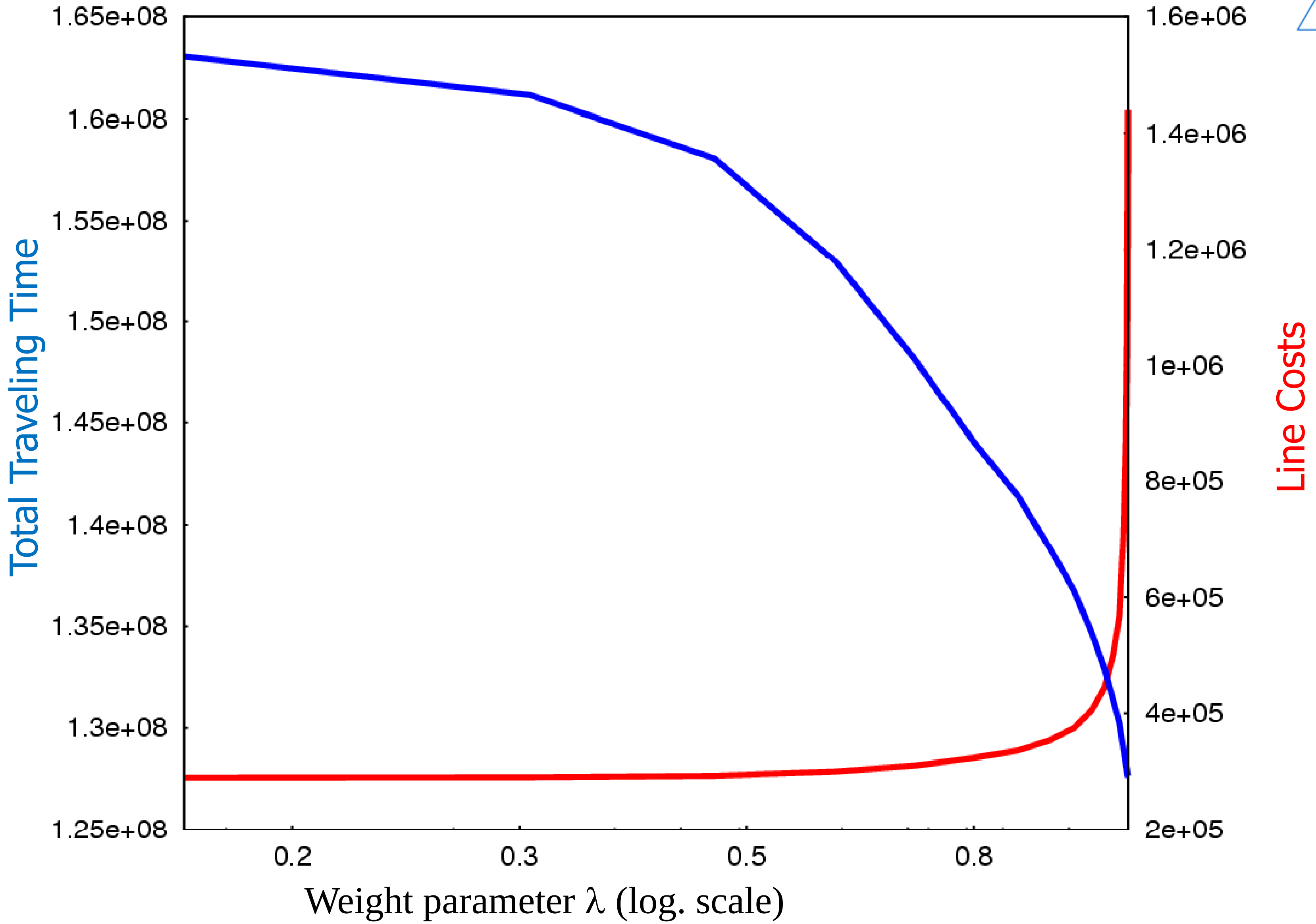
current gap: 0.0474%

Contents

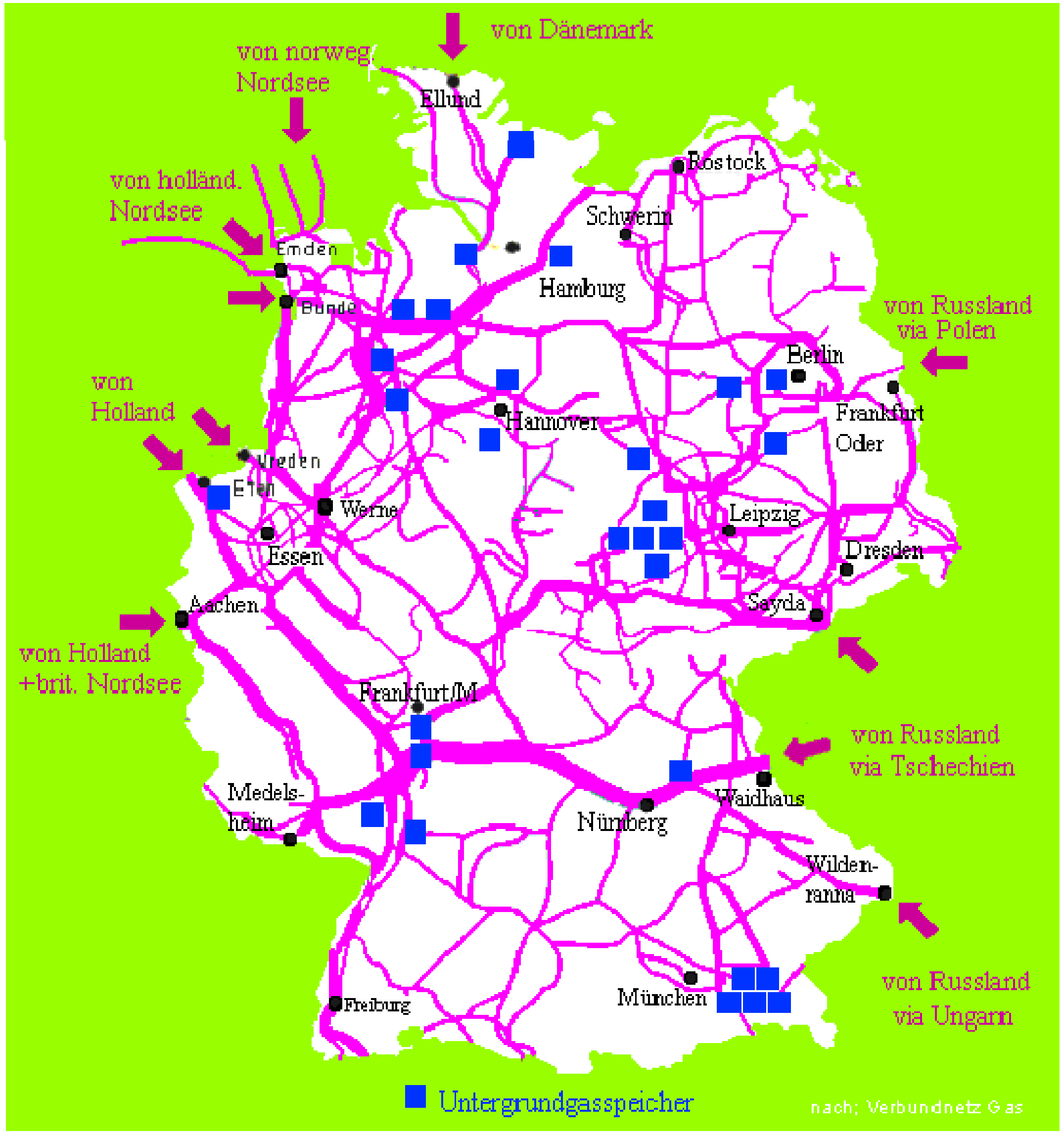
Final comments

- Theoretical versus experimental mathematics: algorithm engineering
- Further challenges:
  - Parallelization for simplex algorithm (for super computers)
  - Parallelization for MIP-solver (for super computers)
  - Warm start for barrier method
  - Finding a basis for barrier
  - Coping with a changing computational environment
  - Coping with huge data and new data environments
  - Reproducibility
  - Solving multi-objective LPs and MIPs
  - Solving MINLPs

Multiobjective optimization
Travel Time vs. Line Costs in Potsdam



German backbone gas pipeline system

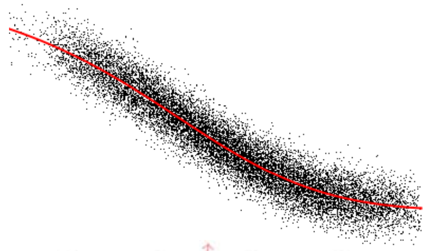


Aspects of Gas Transportation

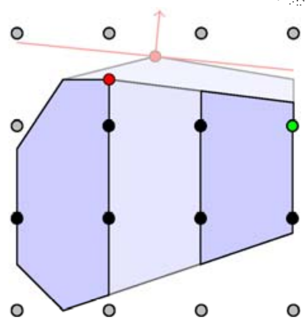
The OGE problem is a:

- Stochastic
- Mixed
- Integer
- Non
- Linear
- Constraint
- Program

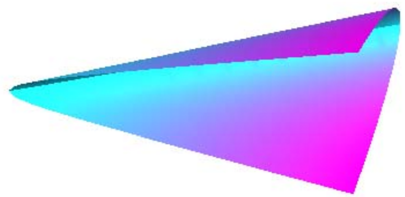
It consists of:



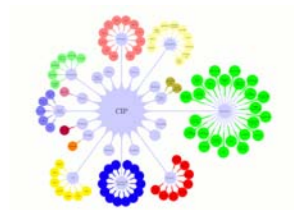
- Stochastic Part



- Mixed Integer Part
- Non-Linear Part



- Constraint Integer Programming Part



Linear and Mixed Integer Optimization: The Solution Methods — Just a Glimpse

Martin Grötschel

CO@W Berlin

Thanks for your attention