Proceedings of the International Conference
In Memoriam Gyula Farkas
August 23–26, 2005, Cluj-Napoca

Proceedings of the International Conference In Memoriam Gyula Farkas

August 23–26, 2005
Cluj-Napoca

Edited by
Z. Kása, G. Kassay, J. Kolumbán

Cluj University Press
2006

This conference was dedicated to the memory of Gyula Farkas (1847–1930), the famous professor in Mathematics and Physics of the University of Kolozsvár/Cluj between 1887–1915.

ORGANIZERS

Babeş-Bolyai University, Cluj-Napoca
Hungarian Operational Research Society, Budapest
Operations Research Committee of the Hungarian Academy of Sciences
Sapientia University, Cluj-Napoca
Gyula Farkas Association for Mathematics and Informatics, Cluj-Napoca
Hungarian Technical Sciences Society of Transylvania, Cluj-Napoca
Transylvanian Museum Association, Cluj-Napoca

SCIENTIFIC ORGANIZING COMMITTEE

Wolfgang BRECKNER (Cluj, Romania), Tibor CSENDES (Szeged, Hungary), Zoltán GÁBOS (Cluj, Romania), József KOLUMBÁN (Cluj, Romania), Sándor KOMLÓSI (Pécs, Hungary), Katalin MARTINÁS (Budapest, Hungary), Petru T. MOCANU (Cluj, Romania), Zsolt PÁLES (Debrecen, Hungary), Tamás RAPCSÁK (Budapest, Hungary), Paul SZILÁGYI (Cluj, Romania), Béla VIZVÁRI (Budapest, Hungary)

LOCAL ORGANIZING COMMITTEE

Gábor KASSAY (Cluj, Romania), Zoltán KÁSA (Cluj, Romania), Gábor KÖLLŐ (Cluj, Romania), Lehel KOVÁCS (Cluj, Romania), Marian MUREŞAN (Cluj, Romania), László NAGY (Cluj, Romania), Ferenc SZENKOVITS (Cluj, Romania)

INVITED SPEAKERS

Katalin MARTINÁS, Boris S. MORDUKHOVICH, András PRÉKOPA, Constantin ZĂLINESCU, Gert WANKA, László NAGY, Zoltán GÁBOS

`http://math.ubbcluj.ro/~farkas`

Contents

Scientific Papers

Linear Inequalities, Duality Theorems and their Financial Applications

András PRÉKOPA*

RUTCOR, Rutgers Center for Operations Research
640 Bartholomew Road, Piscataway, NJ 08854-8003
and
Department of Operations Research
L. Eötvös University, Budapest, Hungary

Abstract. The purpose of this paper is to commemorate the eminent Hungarian mathematician and physicist, Gyula Farkas (1847–1930) on the occasion of the 75th anniversary of his death. First we quote from letters of Farkas to show how he contributed to the development of the world famous Hungarian mathematical scholarship, not only by his own scientific achievements, but also by his organizational activities. Next we show how linear inequality theorems find one step application in financial theory and practice. Finally, we establish a duality relationship in connection with multiobjective optimization and probabilistic constrained stochastic programming.

AMS 2000 subject classifications: 90-03, 90B99, 90C99
Key words and phrases: Gyula Farkas, linear inequality, duality theory

1 About Farkas as a Human Being

It has been shown in a number of publications (see, e.g., Prékopa 1978, 1980, Gábor, 1990, Martinás and Brodszky, 2002) that Farkas contributed fundamental results to both mathematics and physics. Less known are the facts, however, that he was a human being with noble personal qualities and a very efficient organizer. He was a somewhat reserved person but he thought highly of professional and human qualities and helped a lot those whom he considered worthy of it. He did not seek cheap popularity and it was one of the reasons that he was highly respected, inside and outside the university, which he was able to use favorably in university administrative matters. His words had been decisive. The picture that we can form of his character becomes more complete if we quote from his obituary that appeared after his death in the newspaper of "Az Est": "His colleagues, who new him more closely, enthusiastically admired him, his students adored him".

He succeeded to secure employments as professors at the University of Kolozsvár, for Lajos Schlesinger in 1897, Lipót Fejér in 1905, Frigyes Riesz in 1911 and Alfréd Haar in 1912, who later on became of paramount importance in the development of the 20th century Hungarian mathematical school.

* email: prekopa@rutcor.rutgers.edu

Below we quote from Farkas' letters to Lipót Fejér.

After he succeeded to hire Fejér to the University of Kolozsvár, on July 9, 1905:

> "I learn from your letter, dear doctor, that you are spending your time in the family circle... ...please give my regards to your family members, too. Who thinks fondly about our past and future cooperation,
>
> Gyula Farkas"

On October 3, 1911, while he prepares the position for Riesz, at the University of Kolozsvár:

> "The Committee accepted my proposal.. concerning Frigyes Riesz".

On November 4, 1911 he is already working on getting a position for Haar:

> "I am no longer afraid of loosing Haar to abroad, once he comes home".

2 Financial Applications of Linear Inequality Theorems

The Capital Asset Pricing Model (CAPM) establishes a regression relationship between the rate of return R_i on an asset i, and the market return R_M which can be defined, e.g., by the use of the Dow Jones index. It is given by

$$R_i = a_i + b_i R_M + \epsilon_i, \quad i = 1, \dots, m,$$

where ϵ_i, $i = 1, \dots, m$ are independent random variables with $E(\epsilon_i) = 0$, $i = 1, \dots, m$. In this equation the values R_i, R_M are also random variables, R_M is independent of ϵ_i, $i = 1, \dots, m$. The model was formulated by Treynor (1961) and Sharpe (1963, 1964).

Arbitrage pricing theory (APT), formulated by Ross in 1976, is more general than CAPM because it assumes that the rate of return on asset i is a linear function of more than one factors. The basic equation used by APT is the regression equation

$$R_i = a_i + b_{i1} X_1 + \ldots + b_{in} X_n + \epsilon_i, \tag{1}$$

where ϵ_i, $i = 1, \dots, m$ are independent random variables; X_i, $i = 1, \dots, m$ are also random variables, independent of ϵ_i, $i = 1, \dots, m$, $E(\epsilon_i) = E(X_i) = 0$, $i = 1, \dots, m$, and $a_i, b_{i1}, \dots, b_{in}$, $i = 1, \dots, m$ are constants.

We would like to check, if it is possible to create a portfolio out of the assets $i = 1, \dots, m$ in such a way that we buy or sell an amount $|w_i|$ of asset $i = 1, \dots, m$ and

$$\sum_{i=1}^{m} w_i \le 0, \quad \sum_{i=1}^{m} R_i w_i > 0. \tag{2}$$

If such $w_1, \ldots, w_m$ numbers exist, then arbitrage exists because without any positive investment we obtain positive return. The possibility of arbitrage is excluded if for any $w_i, i = 1, \ldots, m$ for which we have (approximately) $\sum_{i=1}^{m} w_i \epsilon_i = 0$, the following condition holds: the relations

$$\begin{aligned} \sum_{i=1}^{m} w_i &\leq 0 \\ \sum_{i=1}^{m} w_i b_{ij} &= 0, \quad j = 1, \ldots, n \end{aligned} \tag{3}$$

imply that

$$\sum_{i=1}^{m} w_i a_i \leq 0. \tag{4}$$

Note that $a_i = E(R_i),\ i = 1, \ldots, m.$

By Farkas' theorem (1901) the above relationship guarantees the existence of real numbers $\lambda_0 \geq 0$, $\lambda_1, \ldots, \lambda_m$ such that

$$E(R_i) = \lambda_0 + \lambda_1 b_{i1} + \ldots + \lambda_n b_{in}, \qquad i = 1, \ldots, m. \tag{5}$$

The number λ_0 belongs to the first relation in (3), and can be identified as the risk free return R_f. The $\lambda_1, \ldots, \lambda_n$ are pricing the sensitivity coefficients in (1). They are also called risk premiums. Note that the inference from (3) and (4) to (5) is mathematically exact but before we came to that we had assumed that the multipliers $w_1, \ldots, w_m$ eliminate the randomness from the portfolio, i.e., $\sum_{i=1}^{m} w_i \epsilon_i = 0$.

What is today called the main arbitrage pricing theorem is formulated in a different way. We look at m securities, numbered by $1, \ldots, m$, and assume that the world can be in n different states. Security i produces a payoff b_{ij} if the state of the world is j. We form the payoff matrix

$$B = \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & & \vdots \\ b_{m1} & b_{m2} & \cdots & b_{mn} \end{pmatrix}.$$

We also assume that, in addition to the matrix B, we are given the security prices which are the components of the vector $q \in R^m$.

A portfolio is a vector $\theta \in R^m$ with real components. Its market value is $q^T\theta$. We say that θ is an arbitrage portfolio, or simply arbitrage, if at least one of the following two conditions is satisfied

$$C_1 : q^T\theta \leq 0 \text{ and } B^T\theta \geq 0,\ B^T\theta \neq 0$$

$$C_2 : q^T\theta < 0 \text{ and } B^T\theta \geq 0.$$

If none of C_1, C_2 holds, then we say that there is no arbitrage. In other terms, if $C_1 \vee C_2$ holds then there is arbitrage and if $\overline{C_1 \vee C_2} = \bar{C_1} \wedge \bar{C_2}$ holds, then there is no arbitrage.

The matrix B^T has size $n \times m$. The set of vectors

$$Q = \left\{ \begin{pmatrix} B^T \\ -q \end{pmatrix} y \,\middle|\, y \in R^m \right\}$$

is a finitely generated cone, i.e., a convex polyhedral cone. The no arbitrage condition can be stated as: the only intersection of Q and the nonnegative orthant $\{x \,|\, x \in R^{n+1}, x \geq 0\}$ is the zero vector.

A vector $x = (x_1, \ldots, x_n)$ is called a state-price vector if $x_i > 0$, $i = 1, \ldots, n$ and $Bx = q$. The following theorem holds true [3].

Theorem 1 *There is no arbitrage if and only if there exists a state-price vector.*

The proof can be based on the following

Theorem 2 (Stiemke 1915) *There exists x with all positive components satisfying $Ax = 0$, if and only if*

$$A^T y \geq 0 \textit{ implies } A^T y = 0.$$

Proof of the arbitrage theorem. We rem that there exists an x with all positive components satisfying $Bx = q$ (a state-price vector) if and only if there exists a vector $(x_1, \ldots, x_{n+1})$ with all positive components such that

$$(B, -q) \begin{pmatrix} x_1 \\ \vdots \\ x_{n+1} \end{pmatrix} = 0.$$

Stiemke's theorem tells us that it happens if and only if

$$\begin{pmatrix} B^T \\ -q^T \end{pmatrix} y \geq 0 \text{ implies } \begin{pmatrix} B^T \\ -q^T \end{pmatrix} y = 0.$$

This is, however, the necessary and sufficient condition for the no arbitrage, as we have noted above. $\square$

Remark 1 A state-price vector exists if and only if the following LP has positive optimum value:

$$\begin{aligned} &\max v \\ &\text{subject to} \\ &Bx = q \\ &v - x_i \leq 0, \quad i = 1, \ldots, n \\ &v \geq 0. \end{aligned}$$

The following is a probabilistic type arbitrage theorem in which the matrix B has an interpretation different from what we gave to it in the preceding theorem.

Suppose there are m wagers and n possible outcomes of a random experiment (game). If the amount x_i is bet on wager i and the outcome of the experiment is the jth one, then the return is

$$x_i b_{ij}, \quad i = 1, \ldots, m; \quad j = 1, \ldots, n,$$

where x_i can be positive, negative or zero. The total return is, for a fixed j,

$$\sum_{i=1}^{m} x_i b_{ij}, \quad j = 1, \ldots, n.$$

If the vector $p = (p_1, \ldots, p_n)$ has the property that

$$\sum_{j=1}^{n} p_j = 1, \quad p_j \geq 0, \quad j = 1, \ldots, n$$

then we call it a probability vector.

Let $p_1, \ldots, p_n$ be the probabilities of the outcomes of the experiment. Then

$$\sum_{j=1}^{n} b_{ij} p_j, \quad i = 1, \ldots, m$$

is the expected return of a unit bet on wager i.

The following holds true [15].

Theorem 3 *Exactly one of the following two assertions is true.*

(1) *There exists an* $x = (x_1, \ldots, x_m)^T$ *such that*

$$\sum_{i=1}^{m} x_i b_{ij} > 0, \quad j = 1, \ldots, n.$$

(2) *There exists a probability vector* p *such that*

$$\sum_{i=1}^{n} b_{ij} p_j = 0, \quad i = 1, \ldots, m.$$

In other words: either there exits a betting scheme $x = (x_1, \ldots, x_m)^T$ that leads to a sure win, or there exists a probability vector p such that the expected return on each wager is zero.

Proof. The proof is an immediate consequence of the following

Theorem 4 (Gordan, 1873) *There exists a vector* $x \geq 0$, $x \neq 0$ *satisfying* $Ax = 0$ *if and only if there is no* y *such that all components of* $A^T y$ *are positive.*

To determine that in a special situation (1) or (2) holds, we look at the linear programming problem

$$\begin{aligned}
&\max v\\
&\text{subject to}\\
&\sum_{i=1}^{m} x_i b_{ij} \geq v\\
&-1 \leq x_i \leq 1, \quad i = 1, \ldots, m.
\end{aligned}$$

If the optimum value is positive, then case (1) holds and simultaneously we obtain optimal (positive, negative or zero) betting proportions.

For more details about arbitrage pricing theory see Duffie (1996).

3 On the Duality Theorem of Linear Programming

The optimization literature attributes to the first exact proof of the duality theorem of linear programming to Gale, Kuhn and Tucker (1951). There are, however, important antecedants of the theorem, one of them is the theorem of Farkas on linear inequalities. In this short note, however, we pay attention to von Neumann's role in the development of the theorem.

George Dantzig, who passed away in 2005 and was one of the principal architects of linear programming, visited von Neumann in 1947, at the Institute for Advanced Study in Princeton. Von Neumann's book with Morgenstern on game theory was published not long before, so it was quite a good idea from Dantzig to have a discussion with Neumann about the relationship between linear programming and game theory. Dantzig had already grasped the primal-dual relationship of linear programming problems, by analyzing the numerical solutions of LP problems but he had no proof of the duality theorem.

Following the meeting with Dantzig, von Neumann wrote a short note entitled "Discussion of a Maximum Principle" and dated November 15–16, 1947. It was circulated in the Institute of Advanced Study in Princeton. The paper was not intended for publication but it is included in the Collected Works of Neumann (Vol. VI., 89–95) with footnotes by H.W. Kuhn and A.W. Tucker. In that paper Neumann gives a proof for the duality theorem. At a point he invokes the fundamental theorem on linear inequalities. Now, the question is: what theorem he had in mind? If it is Farkas' theorem, then there is a slight gap in the proof, as mentioned by Kuhn and Tucker. If, however, it is the theorem of Haar, that had been proved in 1918, then Neumann's proof is complete. The difference between Farkas' theorem and Haar's theorem is minor. Haar's theorem states similar assertion for non-homogeneous linear inequalities and it can be obtained from Farkas' theorem by a simple homogenization technique. Can we assume that von Neumann's genius at once observed how one can come to the theorem for the non-homogeneous case, without being informed about the existence of Haar's result? The answer would be pure speculation but one is inclined to say yes. Anyway, the fact is that Neumann played an important role in the development of the duality theorem of linear programming. That theorem, on the other hand, can be used to give a very simple proof for Neumann's minimax theorem for two-person zero-sum games.

For later use in this paper we present the primal-dual pair of linear programming problems in its most general form:

Primal problem:

$$\begin{aligned} &\max \{c_1^T x_1 + c_2^T x_2\} \\ &\text{subject to} \\ &A_{11}x_1 + A_{12}x_2 \le b_1 \\ &A_{21}x_1 + A_{22}x_2 = b_2 \\ &\qquad x_1 \ge 0. \end{aligned}$$

Dual problem:

$$\begin{aligned} &\min \{b_1^T y_1 + b_2^T y_2\} \\ &\text{subject to} \\ &A_{11}^T y_1 + A_{21}^T y_2 \ge c_1 \\ &A_{12}^T y_1 + A_{22}^T y_2 = c_2 \\ &\qquad y_1 \ge 0. \end{aligned}$$

The duality theorem states that if one of the two problems has feasible solution and finite optimum, then so does the other one and the optimum values are equal.

4 Programming under Probabilistic Constraint

When we formulate a stochastic programming problem, then in the first step we formulate a deterministic problem that would be our optimization problem if we did not have randomness in it. That problem is called underlying deterministic problem or base problem. After we have identified the random variables in the problem, we observe that it loses its original meaning. Then we reformulate it, by the use of some decision principle. The resulting problem is a stochastic programming problem. Consider the following

Base problem:

$$\begin{aligned} &\min c^T x \\ &\text{subject to} \\ &Ax \geq b \\ &Tx \geq \xi \\ &\ \ x \geq 0, \end{aligned} \tag{6}$$

where A is an $m \times n$, T is an $r \times n$ matrix, x, c, b, ξ are vectors of suitable sizes, ξ is a random variable. The decision variable is x.

Based on this we formulate the

Stochastic programming problem:

$$\begin{aligned} &\min c^T x \\ &\text{subject to} \\ &\quad Ax \geq b \\ &P(Tx \geq \xi) \geq p \\ &\qquad x \geq 0, \end{aligned} \tag{7}$$

where p $(0 < p < 1)$ is a prescribed probability. We call (7) probabilistic constrained problem.

Suppose that ξ has a finite support S and let F be the c.d.f. of ξ.

Definition The point $s \in S$ is a p-level efficient (or briefly a p-efficient) point of S if $F(s) \geq p$ and there is no $y \in S$ such that $y \leq s$, $y \neq s$, $F(y) \geq p$.

Let $\{s_1, \dots, s_N\}$ be the set of p-efficient points. Then problem (7) is equivalent to the following:

$$\begin{aligned} &\min c^T x \\ &\text{subject to} \\ &Ax \geq b \\ &Tx \in \bigcup_{i=1}^{N} (s_i + R_+^r) \\ &\ \ x \geq 0. \end{aligned} \tag{8}$$

Problem (8) is a disjunctive programming problem. A relaxation of it (see, Prékopa, Vizvári, Badics, 1998) is:

$$
\begin{aligned}
&\min c^T x\\
&\text{subject to}\\
&Ax \geq b\\
&Tx - \sum_{i=1}^{N} \lambda_i s_i \geq 0\\
&\sum_{i=1}^{N} \lambda_i = 1\\
&x \geq 0,\ \lambda \geq 0.
\end{aligned}
\tag{9}
$$

The dual of this problem is:

$$
\begin{aligned}
&\max\{v + b^T u\}\\
&\text{subject to}\\
&\quad v - s_i^T z \leq 0, \quad i = 1, \ldots, N\\
&A^T u + T^T z \leq c\\
&\quad u \geq 0,\ z \geq 0.
\end{aligned}
\tag{10}
$$

The decision variables are u, z and v.
An equivalent form of problem (10) is the following:

$$
\begin{aligned}
&\max\left\{\min_{1\leq i\leq N} s_i^T z + b^T u\right\}\\
&\text{subject to}\\
&A^T u + T^T z \leq c\\
&\quad u \geq 0,\ z \geq 0,
\end{aligned}
\tag{11}
$$

where the variable v does not appear.

The $s_1, \ldots, s_N$ p-efficient points that appear in the probabilistic constraint in problem (8), produce a multi-objective optimization problem (11), as its dual.

What we have obtained is that the dual of the relaxation of the probabilistic constrained problem (8) is a multi-objective optimization problem. Now we mention the following

Theorem 5 (Prékopa, Vizvári, Badics, 1998) *If* $\{s_1, \ldots, s_N\}$ *is an antichain, i.e., none of them dominates any other one in the set, then for every* $0 < p < 1$ *there exists a c.d.f. such that* $\{s_1, \ldots, s_N\}$ *is the set of its p-efficient points.*

Now, let us start from problem (11), where we assume that the set $\{s_1, \ldots, s_N\}$ is an antichain. If we reformulate it in the form of problem (10) and take the dual of the latter, then we come to problem (9) that is a relaxation of the probabilistic constrained problem (8), written up (based on the above theorem) with $\{c_1, \ldots, c_N\}$ as the set of p-efficient points of the probability distribution of a random vector ξ.

Thus, not the original probabilistic constrained problem but its relaxation is in a primal-dual relationship with a multi-objective optimization problem.

5 A General, Multi-Objective, Probabilistic Constrained Model and a Ruality Relationship

Consider the following optimization problem:

$$\begin{aligned} &\min \left\{ \max_{1\le i\le M} c_i^T x + q^T y \right\} \\ &\text{subject to} \\ &\quad Ax + B\,y \ge b \\ &P(Tx + Wy \ge \xi) \ge p_0 \\ &\qquad x \ge 0,\ y \ge 0, \end{aligned} \tag{12}$$

where the decision variables are x and y. In the above problem we have simultaneously multiple objective function and probabilistic constraint.

Suppose that ξ is discrete and the set of its p_0-efficient points is $\{s_1, \ldots, s_N\}$. The following problem is equivalent to (12):

$$\begin{aligned} &\min \left\{ \max_{1\le i\le M} c_i^T x + q^T y \right\} \\ &\text{subject to} \\ &Ax + By \ge b \\ &Tx + Wy \in \bigcup_{i=1}^{N} (s_i + R_+^r) \\ &\qquad x \ge 0,\ y \ge 0. \end{aligned} \tag{13}$$

A relaxation of this problem is:

$$\begin{aligned} &\min \left\{ \max_{1\le i\le M} c_i^T x + q^T y \right\} \\ &\text{subject to} \\ &Ax + By \qquad\qquad \ge b \\ &Tx + Wy - \sum_{i=1}^{N} \lambda_i s_i \ge 0 \\ &\qquad\qquad \sum_{i=1}^{N} \lambda_i = 1 \\ &\quad x \ge 0,\ y \ge 0,\ \lambda \ge 0. \end{aligned} \tag{14}$$

Introducing the variable t, (14) can be written in the equivalent form:

$$\begin{aligned}
&\min\{t + q^T y\} \\
&\text{subject to} \\
&t - c_i^T x \geq 0, \quad i = 1, \dots, M \\
&Ax + By \geq b \\
&Tx + Wy - \sum_{i=1}^{N} \lambda_i s_i \geq 0 \\
&x \geq 0,\ y \geq 0,\ \lambda \geq 0.
\end{aligned} \tag{15}$$

The dual of the last problem is:

$$\begin{aligned}
&\max\{v + b^T u\} \\
&\text{subject to} \\
&v - s_i^T z \leq 0, \quad i = 1, \dots, N \\
&B^T u + W^T z \leq q \\
&A^T u + T^T z - \sum_{i=1}^{M} \mu_i c_i \leq 0 \\
&\sum_{i=1}^{M} \mu_i = 1 \\
&u \geq 0,\ z \geq 0,\ \mu \geq 0.
\end{aligned} \tag{16}$$

We assume that the set $\{c_1, \dots, c_M\}$ is an antichain. It follows that $\{-c_1, \dots, -c_M\}$ is also an antichain. It follows from the proof of the Theorem of the former section (see Prékopa, Vizvári, Badics 1998) that if we supplement to it a suitable vector d, the obtained $M + 1$ points $-c_1, \dots, -c_M$, d may be regarded as the support of a random vector η whose p_1-efficient points are $-c_1, \dots, -c_M$, where p_1 is an arbitrarily chosen probability, $0 < p_1 < 1$. If we use this, then we can write (16) in the following equivalent form:

$$\begin{aligned}
&\max\left\{\min_{1 \leq i \leq N} s_i^T z + b^T u\right\} \\
&\text{subject to} \\
&W^T z + B^T u \leq q \\
&P(-T^T z - A^T u \geq \eta) \geq p_1 \\
&z \geq 0,\ u \geq 0.
\end{aligned} \tag{17}$$

Thus, the dual of the convexified problem (16) is of the same type as the original problem (12). In this sense we consider problem (12) and (17) a pair of primal-dual probabilistic constrained multi-objective stochastic programming problems. Below we present side by side the pair of primal-dual problems which are the following

$$\begin{aligned} &\min \left\{ \max_{1\le i\le M} c_i^T x + q^T y \right\} \\ &\text{subject to} \\ &\quad Ax + B\,y \ge b \\ &P(Tx + Wy \ge \xi) \ge p_0 \\ &\qquad x \ge 0,\ \ y \ge 0. \end{aligned} \tag{18}$$

and

$$\begin{aligned} &\max \left\{ \min_{1\le i\le N} s_i^T z + b^T u \right\} \\ &\text{subject to} \\ &\quad W^T z + B^T u \le q \\ &P(-T^T z - A^T u \ge \eta) \ge p_1 \\ &\qquad z \ge 0,\ \ u \ge 0. \end{aligned} \tag{19}$$

For the numerical solutions of the problems (15) and (16) we refer to the papers [2], [12], [13], [14] and [16].

If A, B, W, b, q are 0 matrices and vectors, respectively, then the pair of primal-dual problems is:

$$\begin{aligned} &\min \left(\max_{1\le i\le M} c_i^T x \right) \\ &\text{subject to} \\ &P(Tx \ge \xi) \ge p_0 \\ &\qquad x \ge 0 \end{aligned} \tag{20}$$

$$\begin{aligned} &\max \left(\min_{1\le i\le N} s_i^T z \right) \\ &\text{subject to} \\ &P_1(-T^T z \ge \eta) \ge p_1 \\ &\qquad z \ge 0. \end{aligned} \tag{21}$$

These are discrete variants of Komáromi's primal-dual problems (see Komáromi, 1986).

6 The Case of Continuous Distributions

Let ξ, η be two random vectors with strictly quasi-concave c.d.f.'s F and G, respectively. For any $0 < p_0$, $p_1 < 1$ the sets

$$S = \{w \mid F(w) \ge p_0\}$$
$$C = \{w \mid G(w) \ge p_1\}$$

are convex, compact and

$$\tilde{S} = \{w \mid F(w) = p_0\}$$
$$\tilde{C} = \{w \mid G(w) = p_1\}$$

are their boundary sets, respectively. Then, if we take sequences of discrete subsets of $\tilde{S}$, $\tilde{C}$ that are dense in these sets, respectively, then we obtain a pair of problems that are equivalent to problems (18) and (19), respectively.

The problem equivalent to (18) is:

$$\begin{aligned} &\min \left\{ \max_{c \in \tilde{C}} c^T x + q^T y \right\} \\ &\text{subject to} \\ &Ax + B\,y \geq b \\ &Tx + Wy \in \bigcup_{s \in \tilde{S}} (s + R^r_+) \\ &\quad x \geq 0,\ y \geq 0 \end{aligned} \tag{22}$$

and the problem equivalent to (19) is:

$$\begin{aligned} &\max \left\{ \min_{s \in \tilde{S}} s^T z + b^T u \right\} \\ &\text{subject to} \\ &W^T z + B^T u \leq q \\ &-T^T z - A^T u \in \bigcup_{c \in \tilde{C}} (-c + R^n_+) \\ &\quad z \geq 0,\ u \geq 0. \end{aligned} \tag{23}$$

For the resulting problems (18), (19) (or (22), (23)) we can state a duality theorem of which Komáromi's theorem is a special case.

One financial application of the above multi-objective probabilistic constrained model is the following. Suppose we formulate a cash matching problem, where part of the liabilities are covered by cash flows arising from bonds and other sources, e.g., real estates. The prices of the bonds are known but the real estate building costs may not be known. Suppose that the latters are not random variables, but simply unknown. Then it is reasonable to formulate problem (12). Discretizing, we can solve the primal or the dual problem whichever is more convenient.

References

[1] Dantzig, G. B. (1963). *Linear Programming and Extensions.* Princeton University Press, Princeton, N.J.

[2] Dentcheva, D., A. Prékopa and A. Ruszczyński (2000). Concavity and Efficient Points of Discrete Distributions in Probabilistic Programming. *Math. Prog. Ser. A* **89**, 55–77.

[3] Duffie, D. (1996). *Dynamic Asset Pricing Theory.* Princeton University Press, Princeton, N.J.

[4] Farkas, J. (1901). Theorie der einfachen Ungleichungen. *Journal für die reine und angewandte Mathematik* **124**, 1–24.

[5] Gale, D., H.W. Kuhn and A. W. Tucker (1951). Linear Programming and the Theory of Games. In: *Activity Analysis of Production and Allocation*, Wiley, New York, 317–329.

[6] Haar, A. (1918). A lineáris egyenlőtlenségekről. *Math. és Természettud. Értesítő* **36**, 279–296.

[7] Haar, A. (1924). Über lineare Ungleichungen. *Acta Sci. Math. (Szeged)* **2**, 1–19.

[8] Komáromi, É. (1986). Duality in Probabilistic Constrained Linear Programming. In: *System Modelling and Optimization, Proc. 12th IFIP Conference*, Budapest, Hungary, (A. Prékopa, J. Szelezsán and B. Strazicky, eds.), Lecture Notes in Control and Inf. Sci. Vol. 84, 423–429.

[9] Neumann, von, J. (1928). Zur Theorie der Gesellschaftspiele. *Mathematische Annalen* **100**, 295–320.

[10] Neumann, von, J. (1949). Discussion of a Maximum Principle. In: *Collected Works*, (1961–1963, A. H. Taub, ed.), Pergamon Press.

[11] Prékopa, A. (1973). On Logarithmic Concave Measures and Functions. *Acta Sci. Math. (Szeged)* **34**, 335–343.

[12] Prékopa, A. (1990). Dual Method for a One-Stage Stochastic Programming Problem with Random RHS, Obeying a Discrete Probability Distribution. *Zeitschrift für Operations Research* **34**, 441–461.

[13] Prékopa, A. (1995). *Stochastic Programming.* Kluwer Academic Publishers, Dordrecht.

[14] Prékopa, A., B. Vizvári and T. Badics (1998). Programming under Probabilistic Constraint with Discrete Random Variable. In: *New Trends in Mathematical Programming* (F. Giannessi et al. eds.), Kluwer Academic Publishers, Dordrecht, 235–255.

[15] Ross, S.M. (2003), *An introduction to probability models.* Academic Press, London, 2003.

[16] Vizvári, B. (1987). The Integer Programming Background of a Stochastic Integer Programming Algorithm of Dentcheva–Prékopa–Ruszczyński. *Optimization Methods and Software* **17**, 543–559.

Composed convex programming: duality and Farkas-type results

Radu Ioan BOŢ *, Ioan Bogdan HODREA **, and Gert WANKA * * *

Faculty of Mathematics
Chemnitz University of Technology
D-09107 Chemnitz
Germany

Abstract. It is not hard to prove that many convex optimization problems which are already studied in the literature can be rewritten as a particular instance of the following problem: minimize the sum of a convex function and the composition of a convex and K-increasing function with a K-convex one when the variable varies on a given set. Using a conjugate duality approach we construct the Fenchel-Lagrange dual of this general problem. Moreover, using the connections between the optimal objective values of the primal and the dual problem, a Farkas-type result is proved. It is also shown that some recently obtained Farkas-type results are rediscovered as special cases of our statement.

AMS 2000 subject classifications: 49N15, 90C25, 90C46

Key words and phrases: Conjugate functions, conjugate duality, composed convex functions, Farkas-type results

1 Introduction

Since during the last decades the problems generated by the practical needs turned out to be more and more complex, one of the main problems in optimization is to find some methods and conditions which assure the existence of optimal solution for more and more general problems which encompass as special cases the already studied ones.

The problem treated within this paper consists in minimizing the sum of a convex function and the composition of a convex and K-increasing function with a K-convex one when the variable varies on a given set (K is a closed convex cone). Many optimization problems already treated can be derived as special cases of this general optimization problem; among these special cases we would like to mention only the usual problem of minimizing a convex function regarding geometrical and convex inequality constraints. Because of its generality, the problem had recently drawn the attention of many mathematicians and some new results are to be found in the literature ([1], [8], [10]).

In order to provide duality assertions for the problem we treat, we use the same approach as in [2] and [3]. Thus, using an auxiliary variable, to the primal

* email: radu.bot@mathematik.tu-chemnitz.de
** email: bogdan.hodrea@mathematik.tu-chemnitz.de.
* * * email: gert.wanka@mathematik.tu-chemnitz.de

problem we associate an equivalent one, but whose dual can be easier established. In order to determine its dual, to the new optimization problem the classical Lagrange dual problem is attached. Moreover, as the inner infimum of the Lagrange problem can be considered itself as an optimization problem, its Fenchel dual problem is also determined. The construction of the dual, which is actually what we call the Fenchel-Lagrange dual problem, is in detail described and a constraint qualification which ensures strong duality between the primal problem and its dual is also given. Regarding the Fenchel-Lagrange dual problem, let us mention that more about this type of dual problem can be found in [4], [5], [6], [7], [13].

In [6] and [7] Boţ and Wanka have presented some Farkas-type results for inequality systems involving finitely many convex functions using an approach based on the theory of conjugate duality for convex optimization problems. Within the present paper, using weak and strong duality assertions developed for the problem we treat, these results are extended to a more general one. Moreover, it is shown that some results in the literature arise as special cases of the problem we treat.

The paper is organized as follows. Within the second section some definitions and results needed later are presented. A dual for the optimization problem with composed convex functions and the weak and strong duality assertions are established in the third section. Section 4 contains the main result of the paper. The duality acquired in Section 3 allows us to give a Farkas-type theorem. The last section contains Farkas-type results for some particular instances of the initial one and some recent results are rediscovered as special cases.

2 Notations and preliminaries

For the sake of the completeness some well-known definitions and results are presented in the following. As usual, by $\mathbb{R}^k$ is denoted the k-dimensional real space for any nonnegative integer k. All vectors are considered as column vectors. Any column vector can be transposed to a row vector by an upper index ^{T}. By $x^T y = \sum_{i=1}^k x_i y_i$ is denoted the usual inner product of two vectors $x = (x_1, ..., x_k)^T$ and $y = (y_1, ..., y_k)^T$ in $\mathbb{R}^k$. Considering an arbitrary non-empty closed convex cone $K \subseteq \mathbb{R}^k$, the partial ordering induced by the cone is defined by

$$x \leqq_K y \Leftrightarrow y - x \in K, \quad \forall x, y \in \mathbb{R}^k.$$

Let $\mathbb{R}^k$ to be extended by an element ∞ such that for all $x \in \mathbb{R}^k$ it holds $x \leqq_K \infty$. Regarding the partial ordering induced by the cone K over the set $\mathbb{R}^k$, it is not hard to see that it can be naturally extended to the set $\mathbb{R}^k \cup \{\infty\}$ by taking

$$x \leq_K \infty, \quad \forall x \in \mathbb{R}^k \cup \{\infty\}.$$

Moreover, the addition and the multiplication with a scalar are also natural extended setting

$$\infty + x = x + \infty = \infty \text{ and } t\infty = \infty,$$

for any $x \in \mathbb{R}^k \cup \{\infty\}$ and $t \geq 0$.

To the cone K we can associate its dual cone defined by

$$K^* = \big\{\beta \in \mathbb{R}^k : \beta^T x \geq 0, \forall x \in K\big\}.$$

As any $\beta \in K^*$ is actually a real-valued linear functional $\beta : \mathbb{R}^k \to \mathbb{R}$, we consider its natural extension

$$\beta : \mathbb{R}^k \cup \{\infty\} \to \overline{\mathbb{R}} = \mathbb{R} \cup \{\pm\infty\}, \quad \beta(x) = \begin{cases} \beta^T x, \, x \in \mathbb{R}^k, \\ +\infty, \, x = \infty. \end{cases}$$

Let us consider an arbitrary set $X \subseteq \mathbb{R}^n$. By $\text{ri}(X)$, $\text{co}(X)$ and $\text{cl}(X)$ are denoted the *relative interior*, the *convex hull* and the *closure* of the set X, respectively. Furthermore, the *cone* and the *convex cone* generated by the set X are denoted by $\text{cone}(X) = \bigcup_{\lambda \geq 0} \lambda X$ and, respectively, $\text{coneco}(X) = \bigcup_{\lambda \geq 0} \lambda \, \text{co}(X)$. By $v(P)$ we denote the optimal objective value of an optimization problem (P).

If $X \subseteq \mathbb{R}^n$ is given, we consider the following two functions, the *indicator function*

$$\delta_X : \mathbb{R}^n \to \overline{\mathbb{R}}, \quad \delta_X(x) = \begin{cases} 0, & x \in X, \\ +\infty, & \text{otherwise}, \end{cases}$$

and the *support function*

$$\sigma_X : \mathbb{R}^n \to \overline{\mathbb{R}}, \quad \sigma_X(u) = \sup_{x \in X} u^T x,$$

respectively.

For a given function $f : \mathbb{R}^n \to \overline{\mathbb{R}}$, we denote by $\text{dom}(f) = \big\{x \in \mathbb{R}^n : f(x) < +\infty\big\}$ its *effective domain*, by $\text{epi}(f) = \big\{(x, r) : x \in \mathbb{R}^n, r \in \mathbb{R}, f(x) \leq r\big\}$ its *epigraph*, respectively. The function f is called *proper* if its effective domain is a nonempty set and $f(x) > -\infty$ for all $x \in \mathbb{R}^n$.

We consider also the linear operator

$$\mathcal{T} : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R} \times \mathbb{R}^n, \quad \mathcal{T}\big(x, r\big) = (r, x).$$

When X is a nonempty subset of $\mathbb{R}^n$ we define for the function f the *conjugate relative to the set* X by

$$f_X^* : \mathbb{R}^n \to \overline{\mathbb{R}}, \quad f_X^*(p) = \sup_{x \in X} \big\{p^T x - f(x)\big\}.$$

It is easy to observe that for $X = \mathbb{R}^n$ the conjugate relative to the set X is actually the *(Fenchel-Moreau) conjugate function* of f denoted by f^*. Even more, it is trivial to prove that

$$f_X^* = (f + \delta_X)^* \text{ and } \delta_X^* = \sigma_X.$$

Definition 2.1 The function $g : \mathbb{R}^k \to \overline{\mathbb{R}}$ is called *K-increasing* if for all x and y in $\mathbb{R}^k$ such that $x \leqq_K y$ it holds $g(x) \leq g(y)$.

Definition 2.2 Let the function $h : \mathbb{R}^n \to \mathbb{R}^k \cup \{\infty\}$ be given. The function is called *K-convex* if for all $x, y \in \mathbb{R}^n$ and for all $t \in [0,1]$ one has

$$h\big(tx + (1-t)y\big) \leqq_K th(x) + (1-t)h(y).$$

Definition 2.3 Given the functions $f_1, ..., f_m : \mathbb{R}^n \to \overline{\mathbb{R}}$, we call their *infimal convolution* the function

$$f_1 \Box ... \Box f_m : \mathbb{R}^n \to \overline{\mathbb{R}}, \quad (f_1 \Box ... \Box f_m)(x) = \inf \left\{ \sum_{i=1}^m f_i(x_i) : x = \sum_{i=1}^m x_i \right\}.$$

The following statements close this preliminary section.

Theorem 2.1 (cf. [12]) Let $f_1, ..., f_m : \mathbb{R}^n \to \overline{\mathbb{R}}$ be proper convex functions. If the set $\bigcap_{i=1}^m \operatorname{ri}(\operatorname{dom}(f_i))$ is nonempty, then

$$\left(\sum_{i=1}^m f_i \right)^* (p) = (f_1^* \Box ... \Box f_m^*)(p) = \inf \left\{ \sum_{i=1}^m f_i^*(p_i) : p = \sum_{i=1}^m p_i \right\},$$

and for each $p \in \mathbb{R}^n$ the infimum is attained.

Corollary 2.2 (cf. [3]) Let $f_1, ..., f_m : \mathbb{R}^n \to \overline{\mathbb{R}}$ be proper convex functions. If the set $\bigcap_{i=1}^m \operatorname{ri}(\operatorname{dom}(f_i))$ is nonempty, then

$$\operatorname{epi} \left(\left(\sum_{i=1}^m f_i \right)^* \right) = \sum_{i=1}^m \operatorname{epi}(f_i^*).$$

Proposition 2.3 (cf. [3]) Let $f : \mathbb{R}^k \to \overline{\mathbb{R}}$ be a proper function and $\alpha > 0$ a real number. One has

$$\operatorname{epi}\big((\alpha f)^*\big) = \alpha \operatorname{epi}\big(f^*\big).$$

3 Duality for the general problem

Let $X \subseteq \mathbb{R}^n$ be a nonempty convex set and $K \subseteq \mathbb{R}^k$ a nonempty closed convex cone. Consider the functions $f : \mathbb{R}^n \to \overline{\mathbb{R}}$, $g : \mathbb{R}^k \to \overline{\mathbb{R}}$ and $h : \mathbb{R}^n \to \mathbb{R}^k \cup \{\infty\}$, $h = \big(h_1, ..., h_k\big)^T$, such that f is proper and convex, g is proper, convex and K-increasing and h is K-convex. The function g is extended to the space $\mathbb{R}^k \cup \{\infty\}$ by defining $g(\infty) = +\infty$. Moreover, throughout this section two conditions are imposed. First of all, we assume that

$$X \cap \mathrm{dom}(f) \cap h^{-1}\big(\mathrm{dom}(g)\big) \neq \emptyset, \tag{1}$$

where $h^{-1}\big(\mathrm{dom}(g)\big) = \{x \in \mathbb{R}^n : h(x) \in \mathrm{dom}(g)\}$. The second condition we consider is

$$\mathrm{ri}\left(X \cap h^{-1}(\mathbb{R}^k)\right) \cap \mathrm{ri}\left(\mathrm{dom}(f)\right) \neq \emptyset. \tag{2}$$

As a remark, let us mention that these conditions are independent, although at a first look we are tempted to believe that, if the second relation is fulfilled, then the first relation is fulfilled, too.

The problem we work with is

$$(P) \qquad \inf_{x \in X} \big(f(x) + (g \circ h)(x)\big).$$

Regarding this problem, since the relation (1) is fulfilled, it is trivial to see that that the optimal objective value of the problem (P) fulfills $v(P) < +\infty$. Even more, as the function $g \circ h$ is convex, the problem we treat is actually a convex optimization problem with geometric constraints. In order to give a dual problem for (P) we consider the following convex optimization problem

$$(P') \qquad \inf_{\substack{x \in X, y \in \mathrm{dom}(g),\\ h(x) - y \leqq_K 0}} \big(f(x) + g(y)\big).$$

The connection between (P) and (P') is made by the following result.

Theorem 3.1 For the optimal objective values of (P) and (P') we have $v(P) = v(P')$.

Proof. Consider an arbitrary $x \in X$.

If $x \notin \mathrm{dom}(f) \cap h^{-1}\big(\mathrm{dom}(g)\big)$, either $f(x) = +\infty$ or $(g \circ h)(x) = +\infty$ or both, so that $f(x) + (g \circ h)(x) = +\infty \geq v(P')$.

If $x \in \mathrm{dom}(f) \cap h^{-1}\big(\mathrm{dom}(g)\big)$, take $y = h(x) \in \mathrm{dom}(g)$. Then $y - h(x) = 0 \in K$ and the pair (x, y) is obviously feasible to (P'). Even more, as $f(x) + (g \circ h)(x) = f(x) + g(y)$, this equality is enough to secure $f(x) + (g \circ h)(x) \geq v(P')$.

Taking into consideration the inequalities obtained in the two cases considered above, the inequality

$$v(P) \geq v(P')$$

arises as a simple consequence.

In order to prove the reverse inequality, let us consider an arbitrary pair (x, y) feasible to (P').

Let us assume first that $h(x) = \infty$. This would mean that y must be also equal to ∞ and thus $g(y) = +\infty$. But this contradicts the assumption $y \in \mathrm{dom}(g)$ and therefore $h(x) \in \mathbb{R}^k$.

As $h(x) \leqq_K y$ we have that $g(h(x)) \leq g(y)$, so the inequality $f(x) + g(h(x)) \leq f(x) + g(y)$ is also fulfilled. Even more, we get $v(P) \leq f(x) + g(y)$ and, since this inequality is true for an arbitrary pair (x, y) feasible to (P'), the inequality

$$v(P) \leq v(P')$$

follows at hand. This completes the proof. □

This result allows us to affirm that any dual problem of (P') is automatically a dual problem of (P).

To (P') we associate its Lagrange dual problem with $\beta \in K^*$ as dual variable

$$(D) \qquad \sup_{\beta \in K^*} \inf_{\substack{x \in X, \\ y \in \mathrm{dom}(g)}} \Big\{ f(x) + g(y) + \beta^T \big(h(x) - y\big) \Big\}.$$

Using the definition of the conjugate relative to a set, the inner infimum becomes

$$\begin{aligned} &\inf_{\substack{x \in X, \\ y \in \mathrm{dom}(g)}} \big\{ f(x) + g(y) + \beta^T \big(h(x) - y\big) \big\} \\ &= \inf_{x \in X} \big\{ f(x) + \beta^T h(x) \big\} + \inf_{y \in \mathrm{dom}(g)} \big\{ g(y) - \beta^T y \big\} \\ &= -\sup_{x \in X} \big\{ - f(x) - \beta^T h(x) \big\} - \sup_{y \in \mathrm{dom}(g)} \big\{ \beta^T y - g(y) \big\} \\ &= -\big(f + \beta^T h\big)^*_X(0) - g^*(\beta) \\ &= -g^*(\beta) - \inf_{p \in \mathbb{R}^n} \big\{ f^*(p) + \big(\beta^T h\big)^*_X(-p) \big\}, \end{aligned}$$

and, as relation (2) is accomplished, Theorem 2.1 yields that the last infimum is attained.

Remark. Since $\beta(\infty) = +\infty$ for all $\beta \in K^*$, whenever $h(x) = \infty$ we get $(\beta^T h)(x) = \infty$, for all $\beta \in K^*$, and it is not hard to see that with this condition satisfied we get

$$\mathrm{dom}\left(\beta^T h + \delta_X\right) = X \cap h^{-1}(\mathbb{R}^k), \quad \forall \beta \in K^*.$$

Thus we obtain the following formula for the dual problem to (P') and also (P)

$$(D) \qquad \sup_{\substack{p \in \mathbb{R}^n, \\ \beta \in K^*}} \left\{ -g^*(\beta) - f^*(p) - \left(\beta^T h\right)^*_X(-p) \right\}.$$

As a direct consequence of our construction of (D) we get the following weak duality result.

Theorem 3.2 Between the primal problem (P) and the dual (D) weak duality is always satisfied, i.e. $v(P) \geq v(D)$.

The existent literature contains some examples which prove that strong duality is not always fulfilled (see, for example, [13]). Nevertheless, such a situation can be avoided if we consider the following constraint qualification

$$(CQ) \quad \exists x' \in \mathrm{ri}\left(X \cap h^{-1}(\mathbb{R}^k)\right) \cap \mathrm{ri}\left(\mathrm{dom}(f)\right) : h(x') \in \mathrm{ri}\left(\mathrm{dom}(g)\right) - \mathrm{ri}(K).$$

Theorem 3.3 Assume that $v(P)$ is finite. If (CQ) is fulfilled, then between (P) and (D) strong duality holds, i.e. $v(P) = v(D)$ and the dual problem has an optimal solution.

Proof. We actually prove that strong duality holds between the problems (P') and (D). Using Theorem 3.1 the desired result arises as a direct consequence.

To the problem (P') we associate its Lagrange dual

$$(D) \qquad \sup_{\beta \in K^*} \inf_{\substack{x \in X, \\ y \in \mathrm{dom}(g)}} \left\{ f(x) + g(y) + \beta^T \left(h(x) - y\right) \right\}.$$

As the condition (CQ) is fulfilled and all the involved functions are convex, is is well-known from the existing literature ([1], [12]) that between (P') and (D') strong duality holds, i.e. $v(P') = v(D')$ and there exists a $\overline{\beta} \in K^*$ such that

$$v(P') = \inf_{\substack{x \in X, \\ y \in \mathrm{dom}(g)}} \left\{ f(x) + g(y) + \overline{\beta}^T \left(h(x) - y\right) \right\}.$$

As (CQ) is fulfilled we get using the above calculation

$$\inf_{\substack{x \in X, \\ y \in \mathrm{dom}(g)}} \left\{ f(x) + g(y) + \overline{\beta}^T \left(h(x) - y\right) \right\} = -g^*(\overline{\beta}) - \inf_{p \in \mathbb{R}^n} \left\{ f^*(p) + \left(\overline{\beta}^T h\right)^*_X(-p) \right\}$$

and the infimum in the right-hand side is attained. Therefore there exist $\overline{p} \in \mathbb{R}^n$ and $\overline{\beta} \in K^*$ such that

$$v(P') = -g^*(\overline{\beta}) - f^*(\overline{p}) - \left(\overline{\beta}^T h\right)_X^*(-\overline{p}).$$

Using Theorem 3.1 we obtain $v(P) = v(D)$ and $(\overline{p}, \overline{\beta})$ is an optimal solution for (D). □

4 Farkas-type results via weak and strong duality

Using the results presented within the previous section, the following Farkas-type result can be easily proved.

Theorem 4.1 Suppose that (CQ) holds. Then the following assertions are equivalent:

(i) $x \in X \Rightarrow f(x) + (g \circ h)(x) \geq 0$;
(ii) there exist $p \in \mathbb{R}^n$ and $\beta \in K^*$ such that

$$g^*(\beta) + f^*(p) + \left(\beta^T h\right)_X^*(-p) \leq 0. \tag{3}$$

Proof. "$(i) \Rightarrow (ii)$" The statement (i) implies $v(P) \geq 0$ and, since the assumptions of Theorem 3.3 are fulfilled, strong duality holds, i.e. $v(D) = v(P) \geq 0$ and the dual (D) has an optimal solution. Thus there exist $p \in \mathbb{R}^n$ and $\beta \in K^*$ fulfilling (3).

"$(ii) \Rightarrow (i)$" As we can find some $p \in \mathbb{R}^n$ and $\beta \in K^*$ fulfilling (3), it follows right away that

$$v(D) \geq -g^*(\beta) - f^*(p) - \left(\beta^T h\right)_X^*(-p) \geq 0.$$

Weak duality between (P) and (D) always holds and thus we obtain $v(P) \geq 0$, i.e. (i) is true. □

The previous statement can be reformulated as a theorem of the alternative.

Corollary 4.2 Assume that the hypothesis of Theorem 4.1 is fulfilled. Then either the inequality system

$(I) \quad x \in X, f(x) + (g \circ h)(x) < 0$

has a solution or the system

(II) $g^*(\beta) + f^*(p) + \left(\beta^T h\right)^*_X(-p) \leq 0,$
$p \in \mathbb{R}^n, \beta \in K^*$

has a solution, but never both.

Theorem 4.3 The statement (ii) in Theorem 4.1 is equivalent to

$$(0,0,0) \in \{0\} \times \mathcal{T}\left(\operatorname{epi}(g^*)\right) + \operatorname{epi}(f^*) \times \{0\} + \bigcup_{\beta \in K^*} \left(\operatorname{epi}\left((\beta^T h)^*_X\right) \times \{-\beta\}\right).$$

Proof. "$\Rightarrow$" Since the statement (ii) holds, there exist $p \in \mathbb{R}^n$ and $\beta \in K^*$ such that

$$g^*(\beta) + f^*(p) + \left(\beta^T h\right)^*_X(-p) \leq 0.$$

As $g^*(\beta)$ and $(\beta^T h)^*_X(-p)$ have both finite real values, by definition follows

$$\left(\beta, g^*(\beta)\right) \in \operatorname{epi}(g^*)$$

and

$$\left(-p, (\beta^T h)^*_X(-p)\right) \in \operatorname{epi}\left((\beta^T h)^*_X\right).$$

Thus

$$\left(-p, (\beta^T h)^*_X(-p), -\beta\right) \in \operatorname{epi}\left((\beta^T h)^*_X\right) \times \{-\beta\}$$

and it follows

$$\left(-p, \left(\beta^T h\right)^*(-p), -\beta\right) \in \bigcup_{\beta \in K^*} \left(\operatorname{epi}\left((\beta^T h)^*_X\right) \times \{-\beta\}\right). \quad (4)$$

Taking into consideration the definition of the operator $\mathcal{T}$ introduced in the first section of the paper, the relation

$$\left(0, g^*(\beta), \beta\right) \in \{0\} \times \mathcal{T}(\operatorname{epi}(g^*)) \quad (5)$$

follows at once.

On the other hand the inequality

$$f^*(p) \leq -g^*(\beta) - \left(\beta^T h\right)^*(-p)$$

is also fulfilled, and, as the value in the right-hand side is finite, it holds

$$\left(p, -g^*(\beta) - \left(\beta^T h\right)^*_X(-p)\right) \in \operatorname{epi}\left(f^*\right).$$

This implies

$$\left(p, -g^*(\beta) - \left(\beta^T h\right)^*_X(-p), 0\right) \in \operatorname{epi}\left(f^*\right) \times \{0\}. \quad (6)$$

Combining relations (4), (5) and (6) we get

$$(0,0,0) = \big(0,g^*(\beta),\beta\big)+\big(p,-g^*(\beta)-\big(\beta^T h\big)^*_X(-p),0\big) +\big(- p,\big(\beta^T h\big)^*_X(-p),-\beta\big)$$

$$\in \{0\} \times \mathcal{T}\big(\operatorname{epi}(g^*)\big) + \operatorname{epi}(f^*) \times \{0\} + \bigcup_{\beta\in K^*} \Big(\operatorname{epi}\big((\beta^T h)^*_X\big) \times \{-\beta\}\Big).$$

"$\Leftarrow$" Since

$$(0,0,0) \in \{0\}\times\mathcal{T}\big(\operatorname{epi}(g^*)\big) + \operatorname{epi}(f^*) \times \{0\} + \bigcup_{\beta\in K^*}\Big(\operatorname{epi}\big((\beta^T h)^*_X\big) \times \{-\beta\}\Big),$$

we can find some $p \in \mathbb{R}^n$ and $r \in \mathbb{R}$ such that

$$(p, r, 0) \in \operatorname{epi}(f^*) \times \{0\} \tag{7}$$

and

$$(-p, -r, 0) \in \{0\} \times \mathcal{T}(\operatorname{epi}(g^*)) + \bigcup_{\beta\in K^*} \operatorname{epi}\big((\beta^T h)^*_X\big) \times \{-\beta\}. \tag{8}$$

Using the definition of the epigraph of a function, from relation (7) we acquire directly

$$f^*(p) \le r. \tag{9}$$

By relation (8), there exists a $\beta \in K^*$ such that

$$(-p, -r, 0) \in \{0\} \times \mathcal{T}(\operatorname{epi}(g^*)) + \operatorname{epi}\big((\beta^T h)^*_X\big) \times \{-\beta\}.$$

The definition of the operator $\mathcal{T}$ and the previous relation imply that there exist two real numbers r_1 and r_2 such that $-r = r_1 + r_2$, while the pairs (β, r_1) and $(-p, r_2)$ are in $\operatorname{epi}(g^*)$ and $\operatorname{epi}\big((\beta^T h)^*_X\big)$, respectively. Thus

$$g^*(\beta) + (\beta^T h)^*_X(-p) \le r_1 + r_2 = -r. \tag{10}$$

Combining relations (9) and (10), the desired result is straightforward. □

5 The ordinary problem as a particular case

Let $X \subseteq \mathbb{R}^n$ be a nonempty convex set and $K \subseteq \mathbb{R}^k$ a nonempty closed convex cone. Consider the functions $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ and $h : \mathbb{R}^n \to \mathbb{R}^k \cup \{\infty\}$, $h = \big(h_1, ..., h_k\big)^T$, such that f is proper and convex and h is K-convex.

Take the problem

$$(P_1) \qquad \inf_{\substack{x\in X,\\ h(x)\le_K 0}} f(x)$$

and assume that

$$X \cap \operatorname{dom}(f) \cap h^{-1}(-K) \neq \emptyset.$$

It is not hard to remark that for all $x \in \mathbb{R}^n$ we have

$$h(x) \leq_K 0 \Leftrightarrow \delta_{-K}\big(h(x)\big) = 0 \Leftrightarrow (\delta_{-K} \circ h)(x) = 0.$$

Thus we get

$$v(P_1) = \inf_{x \in X} \big(f(x) + (\delta_{-K} \circ h)(x) \big)$$

and, so, (P_1) can be further written as

$$(P_1) \qquad \inf_{x \in X} \big(f(x) + (\delta_{-K} \circ h)(x) \big).$$

Taking into consideration the results obtained in the previous section (to prove that the function δ_{-K} is K-increasing is trivial), to the problem (P_1) we can associate the following dual problem

$$(D_1) \qquad \sup_{\substack{p \in \mathbb{R}^n, \\ \beta \in K^*}} \Big\{ - (\delta_{-K})^*(\beta) - f^*(p) - \big(\beta^T h\big)_X^*(-p) \Big\}.$$

Even more, it is easy to prove that

$$(\delta_{-K})^*(\beta) = \begin{cases} 0, & \beta \in K^*, \\ +\infty, & \text{otherwise}, \end{cases}$$

so that the dual (D_1) becomes

$$(D_1) \qquad \sup_{\substack{p \in \mathbb{R}^n, \\ \beta \in K^*}} \Big\{ - f^*(p) - \big(\beta^T h\big)_X^*(-p) \Big\}.$$

In order to get strong duality between the problems (P_1) and (D_1), the fulfilling of the following constraint qualification is required

$$(CQ_1)\ \exists x' \in \operatorname{ri}\big(X \cap h^{-1}(\mathbb{R}^k)\big) \cap \operatorname{ri}\big(\operatorname{dom}(f)\big) : h(x') \in \operatorname{ri}\big(\operatorname{dom}(\delta_{-K})\big) - \operatorname{ri}(K).$$

But

$$\operatorname{ri}\big(\operatorname{dom}(\delta_{-K})\big) - \operatorname{ri}(K) = \operatorname{ri}(-K) - \operatorname{ri}(K) = -\operatorname{ri}(K) - \operatorname{ri}(K) = -\operatorname{ri}(K),$$

therefore we acquire

$$(CQ_1) \quad \exists x' \in \operatorname{ri}\big(X \cap h^{-1}(\mathbb{R}^k)\big) \cap \operatorname{ri}\big(\operatorname{dom}(f)\big) : h(x') \in -\operatorname{ri}(K).$$

The following outcomes are easy consequences of the results proved within the previous section.

Theorem 5.1 Suppose that (CQ_1) holds. Then the following assertions are equivalent:

(i) $x \in X,\ h(x) \leq_K 0 \Rightarrow f(x) \geq 0$;
(ii) there exist $p \in \mathbb{R}^n$ and $\beta \in K^*$ such that

$$f^*(p) + \big(\beta^T h\big)^*_X(-p) \leq 0.$$

Corollary 5.2 Assume that the hypothesis of Theorem 5.1 is fulfilled. Then either the inequality system

$$(I) \quad x \in X, h(x) \leq_K 0, f(x) < 0$$

has a solution or the system

$$(II)\ f^*(p) + \big(\beta^T h\big)^*_X(-p) \leq 0, \quad p \in \mathbb{R}^n, \beta \in K^*$$

has a solution, but never both.

Theorem 5.3 The statement (ii) in Theorem 5.1 is equivalent to

$$(0,0) \in \operatorname{epi}(f^*) + \bigcup_{\beta \in K^*} \operatorname{epi}\big((\beta^T h)^*_X\big). \tag{11}$$

Proof. By Theorem 4.3 we know that the statement (ii) in Theorem 5.1 is equivalent to

$$(0,0,0) \in \{0\} \times \mathcal{T}\big(\operatorname{epi}((\delta_{-K})^*)\big) + \operatorname{epi}(f^*) \times \{0\} + \bigcup_{\beta \in K^*} \Big(\operatorname{epi}\big((\beta^T h)^*_X\big) \times \{-\beta\}\Big).$$

Since

$$\operatorname{epi}\big((\delta_{-K})^*\big) = K^* \times [0, +\infty),$$

it is easy to see that the last relation can be equivalently written as

$$(0,0,0) \in \bigcup_{\beta \in K^*} \Big(\{0\} \times [0,+\infty) \times K^* + \operatorname{epi}(f^*) \times \{0\} + \operatorname{epi}\big((\beta^T h)^*_X\big) \times \{-\beta\}\Big).$$

This means that there exists $\beta \in K^*$ such that

$$(0,0) \in \{0\} \times [0, +\infty) + \operatorname{epi}(f^*) + \operatorname{epi}\big((\beta^T h)^*_X\big). \tag{12}$$

Using only the definition of the epigraph of a function it is easy to prove that

$$\{0\} \times [0, +\infty) + \operatorname{epi}(f^*) = \operatorname{epi}(f^*).$$

Therefore, by (12),

$$(0,0) \in \operatorname{epi}(f^*) + \bigcup_{\beta \in K^*} \operatorname{epi}\big((\beta^T h)^*_X\big),$$

and the proof is complete. □

Let us consider now $h : \mathbb{R}^n \to \mathbb{R}^k$ and $K = \mathbb{R}^k_+$. The constraint qualification (CQ_1) becomes in this case

$$(CQ_1') \quad \exists x' \in \mathrm{ri}(X) \cap \mathrm{ri}\big(\mathrm{dom}(f)\big) : h(x') \in -\mathrm{ri}\big(\mathbb{R}^k_+\big),$$

which is actually the Slater constraint qualification

$$(CQ_1') \quad \exists x' \in \mathrm{ri}(X) \cap \mathrm{ri}\big(\mathrm{dom}(f)\big) : h(x') < 0.$$

As $\mathrm{ri}(X) \neq \emptyset$, the following equalities can be easily proved (cf. [3], [6])

$$\bigcup_{\beta \in K^*} \mathrm{epi}\,\big((\beta^T h)^*_X\big) = \bigcup_{\beta \geqq 0} \mathrm{epi}\,\big((\beta^T h)^*_X\big) = \mathrm{coneco}\left(\bigcup_{i=1}^{k} \mathrm{epi}(h_i^*)\right) + \mathrm{epi}(\sigma_X).$$

Then the following results are easy consequences of Theorem 5.1 and Theorem 5.3.

Theorem 5.4 Suppose that (CQ_1') holds. Then the following assertions are equivalent:

(i) $x \in X,\ h(x) \leqq 0 \Rightarrow f(x) \geq 0$;
(ii) there exist $p \in \mathbb{R}^n$ and $\beta \geqq 0$ such that

$$f^*(p) + \big(\beta^T h\big)^*_X(-p) \leq 0.$$

Theorem 5.5 The statement (ii) in Theorem 5.4 is equivalent with

$$(0,0) \in \mathrm{epi}(f^*) + \mathrm{coneco}\left(\bigcup_{i=1}^{k} \mathrm{epi}(h_i^*)\right) + \mathrm{epi}(\sigma_X).$$

As a last remark, let us mention that the last two theorems were obtained by Boţ and Wanka in [6], as a generalization of some results due to Jeyakumar ([9]).

6 Conclusions

Within the current paper we deal with conjugate duality and Farkas-type results in composed convex programming. The approach we use is based on conjugate duality for an optimization problem consisting in minimizing the sum between a convex function and the precomposition of an K-increasing and convex function with a K-convex vector function, where K is a closed convex cone. The result we present generalizes some Farkas-type results presented by Boţ and Wanka in [6] and by Jeyakumar in [9]. Moreover, the existing connections between the Farkas-type results and the theorems of the alternative and, respectively, the theory of duality are emphasized once more.

References

[1] Boţ, R.I., Grad, S.-M., Wanka, G., *A new constraint qualification for the formula of the subdifferential of composed convex functions in infinite dimensional spaces*, Preprint 2005–11, Chemnitz University of Technology.

[2] Boţ, R.I., Grad, S.-M., Wanka, G., *A new constraint qualification and conjugate duality for composed convex optimization problems*, Preprint 2004–15, Chemnitz University of Technology.

[3] Boţ, R.I., Hodrea, I.B., Wanka, G., Farkas-type results for inequality systems with composed convex optimization functions via conjugate duality, *Journal of Mathematical Analysis and Applications*, **322**, 1 (2006) 316–328.

[4] Boţ, R.I., Kassay, G., Wanka, G., Strong duality for generalized convex optimization problems, *Journal of Optimization Theory and Applications* 127(1), 2005, pp. 45–70.

[5] Boţ, R.I., Wanka, G., Duality for composed convex functions with applications in location theory, in: W. Habenicht, B. Scheubrein, R. Scheubrein (eds.), *Multi-Criteria- und Fuzzy-Systeme in Theorie und Praxis*, Deutscher Universitäts-Verlag, Wiesbaden, 2003, pp. 1–18.

[6] Boţ, R.I., Wanka, G., Farkas-type results with conjugate functions, *SIAM Journal on Optimization*, 15(2), 2005, pp. 540–554.

[7] Boţ, R.I., Wanka, G., Farkas-type results for max-functions and applications, to appear in *Positivity.*

[8] Combari, C., Laghdir, M., Thibault, L., Sous-différentiels de fonctions convexes composées, *Annales des Sciences Mathématiques du Québec*, 18(2), 1994, pp. 119–148.

[9] Jeyakumar, V., Characterizing set containments involving infinite convex constraints and reverse-convex constraints, *SIAM Journal of Optimization*, 13(4), 2003 pp. 947–959.

[10] Laghdir, M., Volle, M., A general formula for the horizon function of a convex composite function, *Archiv der Mathematik*, 73(4), 1999, pp. 291–302.

[11] Lemaire, B., *Application of a subdifferential of a convex composite functional to optimal control in variational inequalities*, Lecture Notes in Economics and Mathematical Systems 255, 1985, Springer Verlag, Berlin, pp. 103–117.

[12] Rockafellar, R.T., *Convex analysis*, Princeton University Press, Princeton, 1970.

[13] Wanka, G., Boţ, R.I., On the relations between different dual problems in convex mathematical programming, *in:* P. Chamoni, R. Leisten, A. Martin, J. Minnermann and H. Stadtler (eds.), *Operations Research Proceedings 2001*, Springer Verlag, Berlin, 2002, pp. 255–262.

Brézis–Haraux-type approximation of the range of a monotone operator composed with a linear mapping

Radu Ioan BOŢ*, Sorin-Mihai GRAD** and Gert WANKA***

Faculty of Mathematics
Chemnitz University of Technology
D-09107 Chemnitz
Germany

Abstract. We give a Brézis–Haraux-type approximation of the range of the monotone operator $T_A = A^* \circ T \circ A$ when A is a linear continuous mapping between two Banach spaces and T is a maximal monotone operator. Then we specialize the result for a Brézis–Haraux-type approximation of the range of the subdifferential of the precomposition to A of a proper convex lower semicontinuous function defined on a Banach space, which is proven to hold under a weak sufficient condition. This extends and corrects some older results due to Riahi and Chbani that consist in the approximation of the range of the sum of the subdifferentials of two proper convex lower semicontinuous functions.

AMS 2000 subject classifications: 47H05, 46N10, 42A50

Key words and phrases: Monotone operator, range of an operator, subdifferential, Brézis–Haraux approximation

1 Introduction

Given two monotone operators, the sum of their ranges is usually larger than the range of their sum, but there are some situations where these sets are almost equal, i.e. their interiors and closures coincide. Brézis and Haraux ([7], [8]) pioneered the research on this subject giving some conditions that assured the mentioned result in Hilbert spaces. Since then the problem of finding conditions under which the sum of the ranges of two monotone operators is almost equal to the range of their sum is known as the *Brézis–Haraux approximation* problem and the original result has been extended in several directions. Reich ([19]), Chu ([12], [13]) and Simons ([23]) treated the problem in reflexive Banach spaces and Chbani and Riahi ([11]) and Riahi ([20]) in Banach spaces, while Pennanen ([17]), working in reflexive Banach spaces, extended the result from sums of monotone operators to monotone composite mappings of the form $A^* \circ T \circ A$ where A is a linear continuous mapping and T is a monotone operator.

The Brézis–Haraux approximation and its extensions are interesting not only for the results themselves, but also for their many applications. We mention

* email: radu.bot@mathematik.tu-chemnitz.de
** email: sorin-mihai.grad@mathematik.tu-chemnitz.de
*** email: gert.wanka@mathematik.tu-chemnitz.de

here some of them, namely in variational inequality problems ([1]), Hammerstein equations and Neumann problem ([7], [8]), generalized equations ([16]), Kruzkov's solutions of the Burger–Carleman's system ([10]), projection algorithms ([1]), Bregman algorithms ([3]), Fenchel–Rockafellar–Moreau duality model ([16], [17]), optimization problems, Hammerstein differential inclusions and complementarity problems ([11]), and the list is far from being complete.

Within this paper we give a Brézis–Haraux-type approximation statement for $A^* \circ T \circ A$ in Banach spaces. Then we specialize the result to approximate the range of $\partial(f \circ A)$, where f is a proper convex lower semicontinuous function defined on the image space of A with extended real values, generalizing and correcting the result given in [11] and [20] for the sum of the subdifferentials of two proper convex lower semicontinuous functions which arises as special case. Moreover, the regularity condition we impose is weaker than the one considered in the mentioned papers in order to obtain the result. Finally we give two applications, one in optimization and the other to a complementarity problem.

The paper is structured as follows. The next section contains necessary preliminaries, notions and results used later, then we deal with the Brézis–Haraux-type approximation for $A^* \circ T \circ A$. Section 4 deals with the mentioned Brézis–Haraux-type approximations for $\partial(f \circ A)$ and its special case concerning the range of the sum of the subdifferentials of two proper convex lower semicontinuous functions, and it is followed by two applications. An ample list of references closes the paper.

2 Preliminaries

In order to make the paper self - contained we introduce here the context we work within and we recall the necessary notions and results. Let X and Y be two locally convex spaces, unless otherwise specified, and their continuous dual spaces X^* and Y^*, endowed with the weak* topologies $w(X^*, X)$ and $w(Y^*, Y)$, respectively. By $\langle x^*, x\rangle$ we denote the value of the linear continuous functional $x^* \in X^*$ at $x \in X$. Given a subset M of X, we denote by $\operatorname{int}(M)$ and $\operatorname{cl}(M)$ its *interior*, respectively its *closure* in the corresponding topology. We call it *closed regarding the subspace* $Z \subseteq X$ if $M \cap Z = \operatorname{cl}(M) \cap Z$ and we have its *indicator* function $\delta_M : X \to \overline{\mathbb{R}} = \mathbb{R} \cup \{\pm\infty\}$, defined by

$$\delta_M(x) = \begin{cases} 0, & \text{if } x \in M, \\ +\infty, & \text{otherwise.} \end{cases}$$

For a function $f : X \to \overline{\mathbb{R}}$, we have

- the *domain*: $\operatorname{dom}(f) = \{x \in X : f(x) < +\infty\}$,
- the *epigraph*: $\operatorname{epi}(f) = \{(x, r) \in X \times \mathbb{R} : f(x) \le r\}$,
- the *conjugate*: $f^* : X^* \to \overline{\mathbb{R}}$ given by $f^*(x^*) = \sup\{\langle x^*, x\rangle - f(x) : x \in X\}$,

- the *subdifferential* of f at $x \in X$ where $f(x) \in \mathbb{R}$: $\partial f(x) = \{x^* \in X^* : f(y) - f(x) \geq \langle x^*, y - x\rangle \ \forall y \in X\}$,
- f is *proper*: $f(x) > -\infty \ \forall x \in X$ and $\operatorname{dom}(f) \neq \emptyset$.

When $f, g : X \to \overline{\mathbb{R}}$ are proper functions, their *infimal convolution* is defined by

$$f \Box g : X \to \overline{\mathbb{R}}, \ f \Box g(a) = \inf\{f(x) + g(a - x) : x \in X\}.$$

For $f : X \to \overline{\mathbb{R}}$ and $g : Y \to \overline{\mathbb{R}}$, we define the *product function*

$$(f \times g) : X \times Y \to \overline{\mathbb{R}} \times \overline{\mathbb{R}}, \ (f \times g)(x, y) = (f(x), g(y)) \ \forall (x, y) \in X \times Y.$$

Given a linear continuous mapping $A : X \to Y$, its *adjoint* is

$$A^* : Y^* \to X^*, \ \langle A^*y^*, x\rangle = \langle y^*, Ax\rangle \ \forall (x, y^*) \in X \times Y^*.$$

For a proper function $f : X \to \overline{\mathbb{R}}$ we recall also the definition of the *marginal* function of f through A as being

$$Af : Y \to \overline{\mathbb{R}}, \ Af(y) = \inf\big\{f(x) : x \in X, Ax = y\big\} \ \forall y \in Y.$$

Consider also the *identity* function on X defined by

$$\mathrm{id}_X : X \to X, \ \mathrm{id}_X(x) = x \ \forall x \in X.$$

Let us mention moreover that we write min (max) instead of inf (sup) when the infimum (supremum) is attained.

Proposition 1. ([6]) *Let $A : X \to Y$ be a linear continuous mapping and $f : Y \to \overline{\mathbb{R}}$ a proper, convex and lower semicontinuous function such that $f \circ A$ is proper. Then*

(i) $A^* \times \mathrm{id}_{\mathbb{R}}(\mathrm{epi}(f^*))$ is closed in the product topology of $(X^*, w(X^*, X)) \times \mathbb{R}$ if and only if for any $x^* \in X^*$ one has

$$(f \circ A)^*(x^*) = \min\{f^*(y^*) : A^*y^* = x^*\}.$$

(ii) If $A^* \times \mathrm{id}_{\mathbb{R}}(\mathrm{epi}(f^*))$ is closed in the product topology of $(X^*, w(X^*, X)) \times \mathbb{R}$, then for any $x \in \mathrm{dom}(f \circ A)$ one has $\partial(f \circ A)(x) = A^*\partial f(Ax)$.

The second part of this section in devoted to monotone operators and some of their properties. From now on we consider, within the whole paper, X and Y Banach spaces. We denote by $\|\cdot\|$ the norm on X, while the one on X^* is $\|\cdot\|_*$.

Definition 1. ([22]) A mapping (generally multivalued) $T : X \to 2^{X^*}$ is called *monotone operator* provided that for any $x, y \in X$ one has

$$\langle y^* - x^*, y - x\rangle \geq 0 \ \text{ whenever } x^* \in T(x) \text{ and } y^* \in T(y).$$

Definition 2. ([22]) For any monotone operator $T : X \to 2^{X^*}$ we have

· its *effective domain* $D(T) = \{x \in X : T(x) \neq \emptyset\}$,
· its *range* $R(T) = \cup\{T(x) : x \in X\}$,
· its *graph* $G(T) = \{(x, x^*) : x \in X, x^* \in T(x)\}$.

Definition 3. ([22]) A monotone operator $T : X \to 2^{X^*}$ is called *maximal* when its graph is not properly included in the graph of any other monotone operator $T' : X \to 2^{X^*}$.

Let τ_1 be the weakest topology on X^{**} which renders continuous the following real functions

$$X^{**} \to \mathbb{R} : x^{**} \mapsto \langle x^{**}, x^* \rangle \ \forall x^* \in X^*,$$
$$X^{**} \to \mathbb{R} : x^{**} \mapsto \|x^{**}\|.$$

The topology τ in $X^{**} \times X^*$ is the product topology of τ_1 and the strong (norm) topology of X^* (cf. [15]).

Definition 4. ([15]) A monotone operator $T : X \to 2^{X^*}$ is called of *dense type* provided that its *closure* operator $\overline{T} : X^{**} \to 2^{X^*}$,

$$G(\overline{T}) = \left\{(x^{**}, x^*) \in X^{**} \times X^* : \exists (x_i, x_i^*)_i \in G(T) \text{ with } (\hat{x}_i, x_i^*) \xrightarrow{\tau} (x^{**}, x^*)\right\}$$

is maximal monotone, where $\hat{y}$ denotes the canonical image of y in X^{**}.

Different to Riahi ([20]) and Chbani and Riahi ([11]), where these operators are called *densely maximal monotone*, respectively *densely monotone*, we decided to name them as Gossez ([15]) did when he introduced them. By Lemme 2.1 in [15], whenever the monotone operator $T : X \to 2^{X^*}$ is of dense type one has $(x^{**}, x^*) \in G(\overline{T})$ if and only if $\langle x^{**} - \hat{y}, x^* - y^* \rangle \geq 0 \ \forall (y, y^*) \in G(T)$.

The monotone operators belonging to the following class are also known as *star* monotone operators or operators of the *type* (BH), being first introduced in [8].

Definition 5. ([13], [17], [20]) A monotone operator $T : X \to 2^{X^*}$ is called 3^* - *monotone* if for all $x^* \in R(T)$ and $x \in D(T)$ there is some $\beta(x^*, x) \in \mathbb{R}$ such that $\inf_{(y,y^*) \in G(T)} \langle x^* - y^*, x - y \rangle \geq \beta(x^*, x)$.

The last collection of monotone operators we introduce consists of so - called *negative - infimum* monotone operators.

Definition 6. ([23], [24]) A monotone operator $T : X \to 2^{X^*}$ is called of *type* (NI) if for all $(x^{**}, x^*) \in X^{**} \times X^*$ one has $\inf_{(y,y^*) \in G(T)} \langle \hat{y} - x^{**}, y^* - x^* \rangle \leq 0$.

Remark 1. The subdifferential of a proper convex lower semicontinuous function on X is a typical example for all these classes of monotone operators. We refer to [15], [17], [18], [20], [21], [23], [24] and [25] for proofs and more on these subjects.

There are some other types of monotone operators, like cyclic monotone, but as they are not relevant for the results within this paper we do not mention them here. Between these classes of monotone operators there are various relations, let us recall the ones necessary for our purposes.

Proposition 2. ([15]) *In reflexive Banach spaces every maximal monotone operator is of dense type and coincides with its closure operator.*

We close the section by recalling an important result which proved to be useful in the present work.

Lemma 1. ([20]) *Given the dense type operator* $T : X \to 2^{X^*}$ *and the non - empty subset* $E \subseteq X^*$ *such that for any* $x^* \in E$ *there is some* $x \in X$ *fulfilling* $\inf_{(y,y^*)\in G(T)} \langle x^* - y^*, x - y\rangle > -\infty$*, one has* $E \subseteq \mathrm{cl}(R(T))$ *and* $\mathrm{int}(E) \subseteq R(\overline{T})$.

3 Brézis–Haraux-type approximation of the range of a monotone operator composed with a linear mapping

We give in this section the main results concerning the so - called Brézis–Haraux-type approximation (cf. [8], [23]) of the range of a composed operator T_A, defined below, respectively of the subdifferential of the precomposition of a linear continuous mapping with a proper convex lower semicontinuous function. Some results related to ours were obtained by Pennanen in [17], but in reflexive spaces, while we work in general Banach spaces.

Consider the monotone operator $T : Y \to 2^{Y^*}$ and the linear continuous mapping $A : X \to Y$. We introduce the composed operator $T_A := A^* \circ T \circ A : X \to 2^{X^*}$. It is known that T_A is a monotone operator and under certain conditions it is maximal monotone (cf. [4], for instance). We show first that it is 3^* monotone when T is 3^* monotone, too.

Proposition 3. *If* $T : Y \to 2^{Y^*}$ *is* 3^* *- monotone and* $A : X \to Y$ *is a linear continuous mapping, then* T_A *is* 3^* *- monotone, too.*

Proof. If $D(T_A) = \emptyset$, then the conclusion arises trivially. Elsewise take $x^* \in R(T_A)$, i.e. there is some $z \in X$ such that $x^* \in A^* \circ T \circ A(z)$. Thus there exists a $z^* \in T \circ A(z)$ satisfying $x^* = A^* z^*$. Clearly, $z^* \in R(T)$. Consider also an $x \in D(T_A)$ and denote $u = Ax \in D(T)$. When $y^* \in T_A(y)$ there is some $t^* \in T \circ A(y)$ such that $y^* = A^* t^*$. We have

$$\begin{aligned}\inf_{(y,y^*)\in G(T_A)} \langle x^* - y^*, x - y\rangle &= \inf_{(y,t^*)\in G(T\circ A)} \langle A^* z^* - A^* t^*, x - y\rangle \\ &= \inf_{(y,t^*)\in G(T\circ A)} \langle z^* - t^*, A(x - y)\rangle \\ &\geq \inf_{(v,t^*)\in G(T)} \langle z^* - t^*, u - v\rangle \geq \beta(z^*, u) \in \mathbb{R},\end{aligned}$$

as T is 3^* - monotone. Therefore, by definition, T_A is 3^* - monotone, too. □

Next we give an auxiliary result needed in order to prove the main statement of the section which comes after it.

Lemma 2. *If $T : Y \to 2^{Y^*}$ is 3^* - monotone and $A : X \to Y$ is a linear continuous mapping such that T_A is of dense type, then*

(i) $A^*(R(T)) \subseteq \mathrm{cl}(R(T_A))$, *and*
(ii) $\mathrm{int}(A^*(R(T))) \subseteq R(\overline{T_A})$.

Proof. The operator T_A being of dense type implies that $D(T_A) \neq \emptyset$, thus $D(T) \neq \emptyset$.

As T is 3^* - monotone, we have for any $s \in D(T)$ and any $s^* \in R(T)$ there is some $\beta(s^*, s) \in \mathbb{R}$ such that $\beta(s^*, s) \leq \inf_{(y,y^*)\in G(T)} \langle s^* - y^*, s - y\rangle$.

Take some $x^* \in A^*(R(T))$, thus there is an $z^* \in R(T)$ such that $x^* = A^* z^*$. As in the proof of Proposition 3, for some $x \in D(T_A)$ there holds

$$\inf_{(y,y^*)\in G(T_A)} \langle x^* - y^*, x - y\rangle > -\infty.$$

Now we can apply Lemma 1 for $E = A^*(R(T))$ and T_A and we obtain exactly (i) and (ii). □

Theorem 1. *If $T : Y \to 2^{Y^*}$ is 3^* - monotone and $A : X \to Y$ is a linear continuous mapping such that T_A is of dense type, then*

(i) $\mathrm{cl}(A^*(R(T))) = \mathrm{cl}(R(T_A))$, *and*
(ii) $\mathrm{int}(R(T_A)) \subseteq \mathrm{int}(A^*(R(T))) \subseteq \mathrm{int}(R(\overline{T_A}))$.

Proof. The operator T_A being of dense type implies that $D(T_A) \neq \emptyset$. Take some $x^* \in R(T_A)$. Then there are some $x \in X$ and $y^* \in T \circ A(x) \subseteq R(T)$ such that $x^* = A^* y^*$. Thus $x^* \in A^*(R(T))$, so $R(T_A) \subseteq A^*(R(T))$ and the same inclusion stands also between the closures, respectively the interiors, of these sets, i.e.

$$\mathrm{cl}(R(T_A)) \subseteq \mathrm{cl}(A^*(R(T))) \quad \text{and} \quad \mathrm{int}(R(T_A)) \subseteq \mathrm{int}(A^*(R(T))). \tag{1}$$

On the other hand, by Lemma 2(i) we get immediately

$$\mathrm{cl}(A^*(R(T))) \subseteq \mathrm{cl}(R(T_A)) \text{ and } \mathrm{int}(A^*(R(T))) \subseteq \mathrm{int}(R(\overline{T_A})). \tag{2}$$

Relations (i) and (ii) follow immediately from (1) and (2). □

Remark 2. The previous statement generalizes Theorem 1 in [20], which can be obtained for $Y = X \times X$, $Ax = (x, x)$ and $T(y, z) = (T_1(y), T_2(z))$. The next assertion extends Corollary 1 in [20] which arises for the same choice of Y, A and T.

Corollary 1. *If X is a reflexive Banach space, $T : Y \to 2^{Y^*}$ is a 3^* - monotone operator and $A : X \to Y$ is a linear continuous mapping such that T_A is maximal monotone, then one has*

$$\mathrm{cl}(A^*(R(T))) = \mathrm{cl}(R(T_A)) \text{ and } \mathrm{int}(R(T_A)) = \mathrm{int}(A^*(R(T))).$$

Proof. As X is reflexive, Proposition 2 yields that T_A is maximal monotone of dense type and $\overline{T_A}$ and T_A coincide. Theorem 1 delivers the conclusion. □

4 The approximation of the range of the subdifferential of a function composed with a linear mapping

We generalize now Corollary 2 in [20] and Corollary 3.2 in [11], providing a Brézis–Haraux-type approximation of the range of the subdifferential of the precomposition of a proper convex lower semicontinuous function with a linear continuous mapping. Moreover we correct the mentioned results which are improved further by considering a weaker constraint qualification under which one can give the Brézis–Haraux-type approximation of the range of the sum of the subdifferentials of two proper convex lower semicontinuous functions. First we give the constraint qualification that guarantees our more general result,

(CQ) $A^* \times \mathrm{id}_{\mathbb{R}}(\mathrm{epi}(f^*))$ is closed in the product topology of $(X^*, w(X^*, X)) \times \mathbb{R}$.

Theorem 2. *Let the proper convex lower semicontinuous function $f : Y \to \overline{\mathbb{R}}$ and the linear continuous mapping $A : X \to Y$ such that $f \circ A$ is proper, and assume (CQ) valid. Then one has*

(i) $\mathrm{cl}(A^*(R(\partial f))) = \mathrm{cl}(R(\partial(f \circ A)))$, *and*
(ii) $\mathrm{int}(R(\partial(f \circ A))) \subseteq \mathrm{int}(A^*(R(\partial f))) \subseteq \mathrm{int}(D(\partial(A^* f^*)))$.

Proof. As $f \circ A$ is proper, convex and lower semicontinuous, by Théoréme 3.1 in [15] we know that $\partial(f \circ A)$ is an operator of dense type, while according to Theorem B in [21] (see also [17], [20]) ∂f is 3^* - monotone.

By Proposition 1(ii) we know that (CQ) implies $A^* \circ \partial f \circ A = \partial(f \circ A)$. Therefore $A^* \circ \partial f \circ A$ is an operator of dense type, too.
Applying Theorem 1 for $T = \partial f$ we get

$$\mathrm{cl}(A^*(R(\partial f))) = \mathrm{cl}(R(A^* \circ \partial f \circ A)) = \mathrm{cl}(R(\partial(f \circ A)))$$

and

$$\mathrm{int}(R(A^* \circ \partial f \circ A)) \subseteq \mathrm{int}(A^*(R(\partial f))) \subseteq \mathrm{int}(R(\overline{A^* \circ \partial f \circ A})).$$

The relation above that involves closures yields (i), while the other becomes

$$\mathrm{int}(R(\partial(f \circ A))) \subseteq \mathrm{int}(A^*(R(\partial f))) \subseteq \mathrm{int}(R(\overline{\partial(f \circ A)})). \tag{3}$$

As from Proposition 1(i) one may deduce that under (CQ) $A^* f^* = (f \circ A)^*$, by Théoréme 3.1 in [15] we get $R(\overline{\partial(f \circ A)}) = D(\partial(f \circ A)^*) = D(\partial(A^* f^*))$. Putting this into (3) we get (ii). □

When one takes $Y = X \times X$, $Ax = (x, x)$ and $f(x, y) = g(x) + h(y)$, for $x, y \in X$, where g and h are functions defined on X with extended - real values, the constraint qualification (CQ) becomes (cf. [6])

(CQ^s) $\mathrm{epi}(g^*) + \mathrm{epi}(h^*)$ is closed in the product topology of $(X^*, w(X^*, X)) \times \mathbb{R}$

and one obtains the following statement.

Corollary 2. (see also [5]) *Let g and h be two proper convex lower semicontinuous functions on the Banach space X with extended real values such that* $\mathrm{dom}(g) \cap \mathrm{dom}(h) \neq \emptyset$. *Assume (CQ^s) satisfied. Then one has*

(i) $\mathrm{cl}(R(\partial g) + R(\partial h)) = \mathrm{cl}(R(\partial(g + h)))$, *and*
(ii) $\mathrm{int}(R(\partial g + \partial h)) \subseteq \mathrm{int}(R(\partial g) + R(\partial h)) \subseteq \mathrm{int}(D(\partial(g^* \Box h^*))) = \mathrm{int}(D(\partial((g + h)^*)))$.

Proof. We apply Theorem 2 and Proposition 1 for $Ax = (x, x)$ and $f(y, z) = g(y) + h(z)$ for any $(y, z) \in Y = X \times X$. One can easily verify that $(f \circ A)(x) = g(x) + h(x)$, $A^*(y^*, z^*) = y^* + z^*$ $\forall (y^*, z^*) \in X^* \times X^*$ and $A^* f^* = g^* \Box h^*$. Moreover, $A^*(R(\partial f)) = A^*(R(\partial g) \times R(\partial h)) = R(\partial g) + R(\partial h)$ and $A^* \times \mathrm{id}_{\mathbb{R}}(\mathrm{epi}(f^*)) = \mathrm{epi}(g^*) + \mathrm{epi}(h^*)$.

By Proposition 1(ii) we have that (CQ^s) yields $\partial(g + h) = \partial g + \partial h$.

Using the remarks above, from Theorem 2 we get $\mathrm{cl}(R(\partial(g+h))) = \mathrm{cl}(R(\partial g) + R(\partial h))$ and

$$\mathrm{int}(R(\partial(g+h))) \subseteq \mathrm{int}(R(\partial g) + R(\partial h)) \subseteq \mathrm{int}\, D(\partial(g^* \Box h^*)) = \mathrm{int}(D(\partial(g+h)^*)),$$

the last equality arising by Proposition 1(ii). □

We remark that this proof is different from the one in [5].

Similar results have been obtained by Riahi in Corollary 2 in [20] and by Chbani and Riahi in Corollary 3.2 in [11], under the constraint qualification

$$(CQ_R) \qquad \underset{t>0}{\cup} t(\mathrm{dom}(g) - \mathrm{dom}(h)) \text{ is a closed linear subspace of } X.$$

In [20] (CQ_R) is said to imply

$$\mathrm{cl}(R(\partial g)+R(\partial h)) = \mathrm{cl}(R(\partial(g+h))) \text{ and } \mathrm{int}(R(\partial g)+R(\partial h)) = \mathrm{int}(D(\partial(g^* \Box h^*))),$$

while according to [11] it yields

$$\mathrm{cl}(R(\partial g)+R(\partial h)) = \mathrm{cl}(R(\partial(g+h))) \text{ and } \mathrm{int}(R(\partial g)+R(\partial h)) = \mathrm{int}(D(\partial(g+h)^*)).$$

We prove that the latter is not always true when (CQ_R) stands. For a proper, convex and lower semicontinuous function $g : X \to \overline{\mathbb{R}}$ (by taking $h \equiv 0$) Riahi's relation would become $\mathrm{int}(R(\partial g)) = \mathrm{int}(D(\partial g^*))$, which is equivalent, by Théoréme 3.1 in [15], to

$$\mathrm{int}(R(\partial g)) = \mathrm{int}(R(\overline{\partial g})). \tag{4}$$

From Théoréme 3.1 in [15] we know that ∂g is a monotone operator of dense type and, from [21], that it is maximal monotone, too. By [24] we know that ∂g is also of type (NI).

By Theorem 20 in ([24]) we get that $\mathrm{int}(R(\overline{\partial g}))$ is convex, so (4) yields that $\mathrm{int}(R(\partial g))$ is convex, too.

Unfortunately this is not always true, as Example 2.21 in [18], originally given by Fitzpatrick, shows. Take $X = c_0$, the space of the real sequences converging to 0, which is a non - reflexive Banach space with the usual norm $\|x\| = \sup_{n\geq 1} |x_n|$ $\forall x = (x_n)_{n\geq 1} \in c_0$, and $g(x) = \|x\| + \|x - e_1\|$, for all $x \in c_0$, where $e_1 = (1, 0, 0, ...) \in c_0$. It is clear that g is proper, convex and continuous on c_0, since $\|\cdot\|$ has these properties. Moreover for any $x \in c_0$ one has $\partial g(x) = \partial\|\cdot\|(x) + \partial\|\cdot - e_1\|(x)$.

The dual space of c_0 is l^1, which consists of all the sequences $y = (y_n)_{n\geq 1}$ such that $\|y\|_* = \sum_{n=1}^{+\infty} |y_n| < +\infty$. Denote by F the set of sequences in l^1 having finitely many non - zero entries and by B^* the closed unit ball in l^1.

It is known that $\|\cdot\|^*(y) = 0$ if $\|y\|_* \leq 1$ and $\|\cdot\|^*(y) = +\infty$ otherwise, which leads to $\partial\|\cdot\|(x) = B^*$ if $x = 0$, $\partial\|\cdot\|(e_1) = \{e_1^*\}$, $\partial\|\cdot\|(-e_1) = \{-e_1^*\}$ and $\partial\|\cdot\|(x) = \{y \in l^1 : \|y\|_* \leq 1,\ \langle y, x\rangle = \|x\|\} \subseteq F$, otherwise, where $e_1^* = (1, 0, 0, ...) \in l^1$. Moreover we have $\partial\|\cdot - e_1\|(x) = \partial\|\cdot\|(x - e_1)$ for any $x \in c_0$. Further one gets $\partial g(0) = -e_1^* + B^*$ and $\partial g(e_1) = e_1^* + B^*$. Otherwise, i.e. if $x \in c_0 \backslash \{0, e_1\}$, $\partial g(x) \subseteq F$. Therefore

$$R(\partial g) \subseteq (-e_1^* + B^*) \cup (e_1^* + B^*) \cup F. \tag{5}$$

Since $\text{int}(R(\partial g))$ includes $\text{int}(B^*) \pm e_1^*$, assuming it convex yields $0 = 1/2(e_1^* - e_1^*) \in \text{int}(R(\partial g))$. Hence there is a neighborhood of 0, say U, completely included in $R(\partial g)$. Take some $\lambda > 0$ sufficiently small such that

$$\nu(\lambda) = \left(0, \frac{\lambda}{2^2}, \frac{\lambda}{2^3}, \frac{\lambda}{2^4}, \dots\right) \in U.$$

Thus $\nu(\lambda) \in R(\partial g)$. One can check that $\|\nu(\lambda) \pm e_1^*\|_* = 1 + \frac{\lambda}{2} > 1$, so, taking into consideration (5), $\nu(\lambda)$ must be in F. It is clear that this does not happen, thus we have obtained a contradiction. Therefore $\text{int}(R(\partial g))$ is not convex, unlike $\text{int}(R(\overline{\partial g}))$. Thus (4) is false and the same happens to the allegations concerning the interior of the sum of the ranges of two subdifferentials in [11] and [20].

Remark 3. As proven in Proposition 3.1 in [9] (see also [6]), (CQ_R) implies (CQ^s), but the converse is not true, as shown by Example 3.1 in [9]. Therefore our Corollary 2 extends, by weakening the constraint qualification, and corrects Corollary 3.2 in [11] and Corollary 2 in [20].

5 Applications

We give in the following two applications of the results we have presented in the previous section. Both of them generalize some earlier statements that are available in [11] under stronger requirements.

5.1 Existence of a solution to an optimization problem

We work within the framework of Corollary 2, i.e. let g and h be two proper convex lower semicontinuous functions on the Banach space X with extended real values such that $\text{dom}(g) \cap \text{dom}(h) \neq \emptyset$.

Theorem 3. *Assume* (CQ^s) *satisfied and moreover that* $0 \in \text{int}(R(\partial g) + R(\partial h))$. *Then there is a neighborhood* V *of 0 in* X^* *such that* $\forall x^* \in V$ *there is an* $\bar{x} \in \text{dom}(g) \cap \text{dom}(h)$ *where*

$$g(\bar{x}) + h(\bar{x}) - \langle x^*, \bar{x}\rangle = \min_{x \in X}[g(x) + h(x) - \langle x^*, x\rangle].$$

Proof. By Corollary 2 we have $\text{int}(R(\partial g) + R(\partial h)) \subseteq \text{int}(D(\partial((g+h)^*)))$, thus $0 \in \text{int}(D(\partial((g+h)^*)))$, i.e. there is a neighborhood V of 0 in X^* such that $V \subseteq D(\partial((g+h)^*))$. Fix some $x^* \in V$. Immediately one gets that there is some $\bar{x} \in \text{dom}(g) \cap \text{dom}(h)$ such that $(g+h)^*(x^*) + ((g+h)^*)^*(\bar{x}) = \langle x^*, \bar{x}\rangle$.

As $g+h$ is a proper convex lower semicontinuous function we have $(g+h)^{**} = ((g+h)^*)^* = g+h$, thus the equality above becomes

$$g(\bar{x}) + h(\bar{x}) - \langle x^*, \bar{x} \rangle = -(g+h)^*(x^*) = -\sup_{x \in X} \{ \langle x^*, x \rangle - g(x) - h(x) \}.$$

This means actually that the conclusion stands. Because of (CQ^s) we know (cf. [6]) that

$$\inf_{x \in X} [g(x) + h(x) - \langle x^*, x \rangle] = \max_{p \in X^*} \{ -g^*(p) - h^*(x^* - p) \},$$

so one may notice that under the assumptions of the problem we obtain something that may be called locally stable total Fenchel duality, i.e. the situation where both problems, the primal on the left-hand side and the dual on the right-hand side, have optimal solutions and their values coincide for small enough linear perturbations of the objective function of the primal problem. Let us notice moreover that as $0 \in V$, for $x^* = 0$ we obtain also the classical Fenchel strong duality statement, but where moreover the primal problem has a solution, too.□

5.2 Existence of a solution to a complementarity problem

Further consider X a reflexive Banach space, let $C \subseteq X$ be a closed convex cone and $S : X \to 2^{X^*}$ a monotone operator. In order to formulate the statement we have to introduce some new notions and to recall a recent result of ours.

To a monotone operator $T : X \to 2^{X^*}$ Fitzpatrick ([14], see also [4]) attached the function

$$\varphi_T : X \times X^* \to \overline{\mathbb{R}}, \ \varphi_T(x, x^*) = \sup \big\{ \langle y^*, x \rangle + \langle x^*, y \rangle - \langle y^*, y \rangle : (y, y^*) \in G(T) \big\}.$$

For any monotone operator T it is quite clear that φ_T is a convex lower semicontinuous function as an affine supremum. Denote also $\Delta_X = \{(x, x) : x \in X\}$.

Theorem 4. ([5]) *Given two maximal monotone operators T_1 and T_2 on X. If the constraint qualification*

$(\widetilde{CQ})$ $\{(x^* + y^*, x, y, r) : \varphi_{T_1}^*(x^*, x) + \varphi_{T_2}^*(y^*, y) \leq r\}$ *is closed regarding the subspace* $X^* \times \Delta_X \times \mathbb{R}$,

is fulfilled then $T_1 + T_2$ is a maximal monotone operator.

Consider the complementarity problem

$$(CP) \qquad \begin{cases} x \in C, \ x^* \in C^*, \\ \langle x^*, x \rangle = 0, \\ x^* \in S(x). \end{cases}$$

and the constraint qualification

$(\overline{CQ})$ $\{(x^* + y^*, x, y, r)\colon (x^*, x, r) \in \operatorname{epi}(\varphi_S^*),\ y \in C,\ y^* \in -C^*\}$ is closed regarding the subspace $X^* \times \Delta_X \times \mathbb{R}$.

Theorem 5. *Suppose that S is simultaneously maximal and 3^* monotone, assume $(\overline{CQ})$ fulfilled and moreover that $0 \in \operatorname{int}(R(S) - C^*)$. Then the complementarity problem (CP) admits a solution.*

Proof. The conjugate function to δ_C and its subdifferential are

$$\delta_C^*(y^*) = \begin{cases} 0, & \text{if } y^* \in -C^*, \\ +\infty, & \text{otherwise.} \end{cases} \quad \text{and } \partial\delta_C(x) = \{y^* \in -C^* : \langle y^*, x\rangle = 0\}\ \forall x \in C.$$

It is easy to notice that $R(\partial\delta_C) \subseteq -C^*$ and $\partial\delta_C(0) = -C^*$, thus $R(\partial\delta_C) = -C^*$.

It is also straightforward to see that finding a solution to (CP) is equivalent to proving the existence of some $x \in C$ such that $0 \in S(x) + \partial\delta_C(x) = (S + \partial\delta_C)(x)$.

In order to apply Corollary 1 we need the maximal monotonicity of $S+\partial\delta_C$. As suggested by Theorem 4 we calculate the Fitzpatrick function attached to $\partial\delta_C$ and its conjugate. We have for some pair $(x, x^*) \in X \times X^*$

$$\begin{aligned} \varphi_{\partial\delta_C}(x, x^*) &= \sup_{(y,y^*)\in G(\partial\delta_C)} \{\langle y^*, x\rangle + \langle x^*, y\rangle - \langle y^*, y\rangle\} \\ &= \sup_{\substack{y\in C, y^*\in -C^*,\\ \langle y^*, y\rangle = 0}} \{\langle y^*, x\rangle + \langle x^*, y\rangle\} \\ &= \begin{cases} 0, & \text{if } x \in C, x^* \in -C^*, \\ +\infty, & \text{otherwise.} \end{cases} \end{aligned}$$

Its conjugate is, for $(z^*, z) \in X^* \times X$,

$$\varphi_{\partial\delta_C}^*(z^*, z) = \sup_{\substack{x\in C,\\ x^*\in -C^*}} \{\langle z^*, x\rangle + \langle x^*, z\rangle\} = \begin{cases} 0, & \text{if } z \in C,\ z^* \in -C^*, \\ +\infty, & \text{otherwise.} \end{cases}$$

It is not difficult to observe now that for $T_1 = S$ and $T_2 = \partial\delta_C$ the constraint qualification $(\widetilde{CQ})$ turns into $(\overline{CQ})$. This leads, by Theorem 4, to the maximal monotonicity of $S + \partial\delta_C$, so by Corollary 1, for the same choice of Y, A and S as in Remark 2, one gets

$$\operatorname{int}(R(S) - C^*) = \operatorname{int}(R(S) + R(\partial\delta_C)) = \operatorname{int}(R(S + \partial\delta_C)),$$

as in this case $T_A = S + \partial\delta_C$ and $A^*R(T) = R(S) + R(\partial\delta_C)$.

From the hypothesis we get $0 \in \operatorname{int}(R(S + \partial\delta_C))$, thus $0 \in R(S + \partial\delta_C)$, i.e. there is some $x \in C$ such that $0 \in S(x) + \partial\delta_C(x) = (S + \partial\delta_C)(x)$. As remarked above, this is equivalent to the fact that (CP) admits a solution. $\square$

References

[1] Auslender, A., Teboulle, M., Lagrangian duality and related multiplier methods for variational inequality problems, *SIAM Journal on Optimization*, No. 10 (2000), 1097-1115.

[2] Bauschke, H.H., The composition of projections onto closed convex sets in Hilbert space is asymptotically regular, *Proceedings of the American Mathematical Society*, No. 131 (2003), 141–146.

[3] Bauschke, H.H., Borwein, J.M.,Combettes, P., Bregman monotone optimization algorithms, SIAM *Journal on Control and Optimization,* No. 42 (2003), 596–636.

[4] Borwein, J.M., Maximal monotonicity via convex analysis, *Journal of Convex Analysis* (to appear).

[5] Boţ, R.I., Grad, S.M., Wanka, G., *Weaker constraint qualifications in maximal monotonicity*, to appear in *Numerical Functional Analysis and Optimization* (2007).

[6] Boţ, R.I., Wanka, G., A weaker regularity condition for subdifferential calculus and Fenchel duality in infinite dimensional spaces, *Nonlinear Analysis: Theory, Methods & Applications* **64** (2006) 2787–2804.

[7] Brézis, H., Quelques propriétés des opérateurs monotones et des semi - groupes non linéaires, Nonlinear operators and the calculus of variations (Summer School, Université Libre de Bruxelles, Brussels, 1975), *Lecture Notes in Mathematics*, No. 543 (1976), Springer, 56–82.

[8] Brézis, H., Haraux, A., Image d'une somme d'opérateurs monotones et applications, *Israel Journal of Mathematics*, No. 23 (1976), 165–186.

[9] Burachik, R.S., Jeyakumar, V., A dual condition for the convex subdifferential sum formula with applications, *Journal of Convex Analysis*, No. 12 (2005), 279–290.

[10] Caselles, V., Klaiany, C., Existence, uniqueness and regularity for Kruzkov's solutions of the Burger–Carleman's system, *Annales Mathématiques de la Faculté des Sciences de Toulouse 5e série*, No. 10 (1989), 93–104.

[11] Chbani, Z., Riahi, H., The range of sums of monotone operators and applications to Hammerstein inclusions and nonlinear complementarity problems, Nonlinear partial differential equations (Fés, 1994), *Pitman Research Notes in Mathematics Series*, No. 343 (1996), Longman, 61–72.

[12] Chu, L. -J., On Brézis–Haraux approximation with applications, *Far East Journal of Mathematical Sciences*, No. 4 (1996), 425–442.

[13] Chu, L. -J., On the sum of monotone operators, *The Michigan Mathematical Journal*, No. 43 (1996), 273–289.

[14] Fitzpatrick, S., Representing monotone operators by convex functions, in: Workshop/Miniconference on Functional Analysis and Optimization (Canberra, 1988), *Proceedings of the Centre for Mathematical Analysis*, No. 20 (1988), Australian National University, Canberra, 59–65.

[15] GOSSEZ, J. -P., Opérateurs monotones non linéaires dans les espaces de Banach non réflexifs, *Journal of Mathematical Analysis and Applications*, No. 34 (1971), 371–395.

[16] PENNANEN, T., Dualization of generalized equations of maximal monotone type, *SIAM Journal on Optimization,* No. 10 (2000), 809–835.

[17] PENNANEN, T., On the range of monotone composite mappings, *Journal of Nonlinear and Convex Analysis*, No. 2 (2001), 193–202.

[18] PHELPS, R.R., Lectures on maximal monotone operators, *Extracta Mathematicae,* No. 12 (1997), 193–230.

[19] REICH, S., The range of sums of accretive and monotone operators, *Journal of Mathematical Analysis and Applications*, No. 68 (1979), 310–317.

[20] RIAHI, H., On the range of the sum of monotone operators in general Banach spaces, *Proceedings of the American Mathematical Society*, No. 124 (1996), 3333–3338.

[21] ROCKAFELLAR, R.T., On the maximal monotonicity of subdifferential mappings, *Pacific Journal of Mathematics*, No. 33 (1970), 209–216.

[22] ROCKAFELLAR, R.T., On the maximality of sums of nonlinear monotone operators, *Transactions of the American Mathematical Society*, No. 149 (1970), 75–88.

[23] SIMONS, S., *Minimax and monotonicity*, Springer, 1998.

[24] SIMONS, S., The range of a monotone operator, *Journal of Mathematical Analysis and Applications*, No. 199 (1996), 176–201.

[25] ZĂLINESCU, C., *Convex analysis in general vector spaces*, World Scientific, 2002.

A Little Theory for the Control of an Assembly Robot Using Farkas Theorem

Zsolt ROBOTKA and Béla VIZVÁRI*

Department of Operations Research
Eötvös Lorá nd University of Budapest
Budapest, Pázmány Péter sétány 1/C
H-1117 Hungary

Abstract. In this paper the move of a robot arm is optimized via Benders decomposition. The problem is modeled by a combinatorial optimization problem. By applying the famous Farkas theorem it can be shown that even the continuous part of the decomposition reserves its original combinatorial nature.

AMS 2000 subject classifications: 93C05, 90C27

Key words and phrases: Farkas theorem, Benders decomposition, assigment problem, assembly robot

1 Introduction

The problem of this paper has been motivated by printed circuit assembly. A good survey on this topic is [2]. The results of the present paper can be applied in all cases when a robot assembles a product and the objective is to minimize the length of the move of the robot arm.

2 Technological arrangement

The task being assembled by the robot is in a fixed position. The components are in a sequence of cells. Each cell contains a component of different type. Each component has a well-defined position on the task where it is to be assembled. The duration of the assembly of a component is an a priori given fixed value. The only possibility to save time, i.e. to accelerate production, is to minimize the total move of the robot arm.

When the assembly of a component is finished, the arm goes for the next component to the appropriate cell from where there it goes to the position of the next component on the task. Hence it follows that the total move of the arm depends on both *(i) the assignment of the components to cells*, and *(ii) the order of the components in which they are assembled.* Therefore the whole problem is the "direct product" of the assignment problem of (i), and the traveling saleman

* email: vizvari@cs.elte.hu

problem (TSP) of (ii). The distances among cells and positions are supposed to be symmetric.

In this paper we shall suppose that the following two assumptions hold:

$$\text{The number of cells is equal to the number of components.} \tag{A1}$$

$$\text{Each component is used only once.} \tag{A2}$$

These assumptions simplify the problem, which still remains difficult enough to be solved.

3 Problem formulation

To describe the problem mathematically the following notations are introduced:

n	the number of cells and components
i	the index of cells
j, k, l	indices of components and positions
d_{ik}	the symmetric distance of cell i and position k
x_{ij}	is 1 if component j is assigned to cell i, otherwise it is 0
y_{kl}	is 1 if component l is assembled immediately after component k

Variables x and y are the decision variables. They must satisfy the following constraints.

Each component is assigned to exactly one cell and vice versa:

$$\sum_{j=1}^{n} x_{ij} \;=\; 1, \quad i = 1, ..., n, \tag{1}$$

and

$$\sum_{i=1}^{n} x_{ij} \;=\; 1, \quad j = 1, ..., n. \tag{2}$$

Each component is assembled exactly once, i.e. the order of the components is a Hamiltonian circuit:

$$\sum_{k=1}^{n} y_{kl} \;=\; 1, \quad l = 1, ..., n, \tag{3}$$

and

$$\sum_{l=1}^{n} y_{kl} \;=\; 1, \quad k = 1, ..., n, \tag{4}$$

and

$$\forall H \subset \{1, 2, ..., n\}, n > \mid H \mid \geq 2 : \quad \sum_{k \in H} \sum_{l \in \overline{H}} y_{kl} \ \geq \ 1. \tag{5}$$

All variables are binary:

$$x_{ij},\, y_{kl} \ \in \ \{\, 0,\, 1\, \}, \quad i, j, k, l = 1, 2, ..., n. \tag{6}$$

The arm moves from position k of the task to cell i only if cell i contains the component of the immediate successor position. Assume that the its index is l. Thus until the next position the length of the move is $d_{ik} + d_{il}$. These terms of the distance function can be selected by the decision variables and the total distance of the move of the arm to be minimized is:

$$\min \sum_{k=1}^{n} \sum_{l=1}^{n} \sum_{i=1}^{n} d_{ik} y_{kl} x_{il} \ + \ \sum_{i=1}^{n} \sum_{l=1}^{n} d_{il} x_{il}. \tag{7}$$

Thus the mathematical problem to be solved is to optimize (7) under the conditions (1-6). This problem formulation has two drawbacks. At first there are exponential many constraints in (5). We shall see that only those of them will be used, which are violated. Secondly, the objective function (7) is nonlinear. It can be linearized by the usual method. New variables, say w_{ikl}'s, are introduced as follows:

$$w_{ikl} \ = \ x_{il} y_{kl} \quad i, k, l = 1, 2, ..., n. \tag{8}$$

If both x_{il} and y_{kl} are zero-one variables then w_{ikl} is zero-one as well. It is well-known that equation (8) is equivalent to the following inequalities

$$w_{ikl} \ \geq \ x_{il} + y_{kl} - 1 \quad i, k, l = 1, 2, ..., n \tag{9}$$

and

$$2w_{ikl} \ \leq \ x_{il} + y_{kl} \quad i, k, l = 1, 2, ..., n \tag{10}$$

assuming that

$$w_{ikl} \ \in \ \{\, 0,\, 1\, \}, \quad i, k, l = 1, 2, ..., n. \tag{11}$$

Thus the new form of the objective function when it is multiplied by (-1) is

$$\max \sum_{i=1}^{n} \sum_{l=1}^{n} (-d_{il}) x_{il} \ + \ \sum_{k=1}^{n} \sum_{l=1}^{n} \sum_{i=1}^{n} (-d_{ik}) w_{ikl}. \tag{12}$$

Consequently, our aim is to optimize (12) under the conditions (1)-(6), and (10)-(11).

For the sake of convenience we need a compact form of the constraints, too. Inequalities (1)-(2) and (3)-(5), respectively, contain only the variables $\mathbf{x}$ and $\mathbf{y}$, respectively. Thus these sets can be written separately. If it is necessary the inequalities are multiplied by (-1). The final form is as follows

$$\begin{array}{rl} \max\ \mathbf{c}^T\mathbf{x} + \mathbf{0}^T\mathbf{y} + \mathbf{f}^T\mathbf{w} & \\ \mathbf{A}_1\mathbf{x} + \mathbf{O}\mathbf{y} + \mathbf{O}\mathbf{w} & = \mathbf{e}_{2n} \\ \mathbf{O}\mathbf{x} + \mathbf{B}_2\mathbf{y} + \mathbf{O}\mathbf{w} & = (\leq)\ \mathbf{b}_2 \\ \mathbf{A}_3\mathbf{x} + \mathbf{B}_3\mathbf{y} + \mathbf{C}_3\mathbf{w} & \leq \mathbf{b}_3 \\ \mathbf{x}, \mathbf{y} \in \{0,1\}^{n^2},\ \ \mathbf{w} \in \{0,1\}^{n^3}, & \end{array} \tag{13}$$

where vectors $\mathbf{c}$, and $\mathbf{f}$ are formed from the distances according to (12) and all the components of the $2n$-dimensional vector $\mathbf{e}_{2n}$ are 1 and finally $\mathbf{O}$ is a zero matrix of appropriate size. When a uniform notation is more convenient, $\mathbf{b}_1$ will be used instead of $\mathbf{e}_{2n}$.

4 Benders decomposition in the general case

The Benders decomposition [1] is summarized in this section as it is the main tool to develop our algorithm. It is very rarely referred in the literature. This section does not contain new results.

The Benders decomposition is actually the dual of the Dantzig-Wolfe decomposition. Here not the constraints but the variables are divided into two parts. The first one represents a linear programming part while the second one is arbitrary. The problem to solve is

$$\begin{array}{l} \max\ \mathbf{c}^T\mathbf{p} + f(\mathbf{r}) \\ \quad \mathbf{A}\mathbf{p} + \mathbf{F}(\mathbf{r}) \leq \mathbf{b} \\ \quad \mathbf{p} \geq \mathbf{0},\ \mathbf{r} \in \mathcal{S}, \end{array} \tag{14}$$

where $\mathbf{c}$, and $\mathbf{p}$ are s-dimensional vectors, $\mathbf{A}$ is a real matrix of size $m \times s$, $f : \mathbf{R}^t \to \mathbf{R}$, and $\mathbf{F} : \mathbf{R}^t \to \mathbf{R}^m$ are arbitrary functions, $\mathcal{S}$ is an arbitrary subset of $\mathbf{R}^t$ and $\mathbf{r}$ is a t-dimensional vector.

For a fixed $\hat{\mathbf{r}}$ Problem (14) becomes the linear programming problem

$$\begin{array}{l} \max\ \mathbf{c}^T\mathbf{p} + f(\hat{\mathbf{r}}) \\ \quad \mathbf{A}\mathbf{p} \leq \mathbf{b} - \mathbf{F}(\hat{\mathbf{r}}) \\ \quad \mathbf{p} \geq \mathbf{0}, \end{array} \tag{15}$$

where the term $f(\hat{\mathbf{r}})$ in the objective function is only an additive constant. The dual of (15) is

$$\begin{array}{l} \min\ (\mathbf{b} - \mathbf{F}(\hat{\mathbf{r}}))^T\mathbf{u} \\ \quad \mathbf{A}^T\mathbf{u} \geq \mathbf{c} \\ \quad \mathbf{u} \geq \mathbf{0}. \end{array} \tag{16}$$

If Problem (16) has no feasible solution then Problem (15) is either unbounded or has no feasible solution for each particular $\hat{\mathbf{r}}$. Hence the original problem has no optimal solution. Therefore in the rest of the paper it is assumed that Problem (16) has at least one feasible solution.

An equivalent form of (14) can be obtained by introducing an objective function variable, say z, and slack variables, say v_0 and $\mathbf{v}$, to obtain equations instead of the inequalities. The new form of (14) is

$$\begin{array}{rl} \max & \mathbf{0}^T\mathbf{p} + \mathbf{0}^T\mathbf{r} + z + 0v_0 + \mathbf{0}^T\mathbf{v} \\ & -\mathbf{c}^T\mathbf{p} - f(\mathbf{r}) + z + v_0 + \mathbf{0}^T\mathbf{v} = 0 \\ & \mathbf{A}\mathbf{p} + \mathbf{F}(\mathbf{r}) + \mathbf{0}z + \mathbf{0}v_0 + \mathbf{v} = \mathbf{b} \\ & \mathbf{p} \geq \mathbf{0},\ \mathbf{r} \in \mathcal{S},\ v_0 \geq 0,\ \mathbf{v} \geq \mathbf{0}. \end{array} \tag{17}$$

The following theorem is an immediate consequence of the Farkas theorem taking into account that the variables in Problem (17) with the possible exception of $\mathbf{r}$, and z are nonnegatives.

Theorem 6 *For a given pair* $(\hat{\mathbf{r}}, \hat{z})$*, where* $\hat{\mathbf{r}} \in \mathcal{S}$ *and* $\hat{z} \in \mathbb{R}$ *there exist vectors* $\hat{\mathbf{p}}$*, and* $\hat{\mathbf{v}}$ *and a number* $\hat{v}_0$ *such that the 5-tuple* $(\hat{\mathbf{p}}, \hat{\mathbf{r}}, \hat{z}, \hat{v}_0, \hat{\mathbf{v}})$ *is a feasible solution of Problem (17) if and only if the inequality*

$$u_0(f(\hat{\mathbf{r}}) - \hat{z}) + \mathbf{u}^T(\mathbf{b} - \mathbf{F}(\hat{\mathbf{r}})) \ \geq\ 0 \tag{18}$$

holds for every real number u_0 *and vector* $\mathbf{u} \in \mathbb{R}^m$ *such that*

$$\mathbf{A}^T\mathbf{u} \ \geq\ \mathbf{c}u_0, \quad u_0 \ \geq\ 0, \quad \mathbf{u} \ \geq\ \mathbf{0}. \tag{19}$$

The set of the $m{+}1$-dimensional vectors $(u_0, \mathbf{u}^T)^T$ satisfying (19) is obviously a pointed and polyhedral cone denoted by $\mathcal{C}$. It is well-known that it is spanned by the finite set of its extremal directions, say $\mathcal{Q}$. If Inequality (18) holds for all elements of $\mathcal{Q}$ then it holds for all of the elements of $\mathcal{C}$. From computational point of view the problem is that the set $\mathcal{Q}$ may have too many elements to explore all of them as an initial step of the algorithm. Therefore a "column generation" type algorithm should be developed, which uses only the elements of $\mathcal{Q}$, that are really required. As it will be seen this type of algorithm is of "row generation" in the case of Benders decomposition.

Furthermore, if $(\hat{u}_0, \hat{\mathbf{u}}) \in \mathcal{Q}$ and $\hat{u}_0 \neq 0$ then without loss of generality we may assume that $\hat{u}_0 = 1$. Hence one can conclude that to test if a given pair $(\hat{\mathbf{r}}, \hat{z})$ is a part of an optimal solution it is enough to solve Problem (16). Let $\hat{\mathbf{u}}^*$ be the optimal solution. Only the following cases may occur:

(i) optimal solution: If an optimal solution of Problem (16) exists and the equation

$$\hat{z} \ =\ (\mathbf{b} - \mathbf{F}(\hat{\mathbf{r}}))^T\hat{\mathbf{u}}^* + f(\hat{\mathbf{r}}) \tag{20}$$

holds then the pair is optimal and the missing part $\hat{\mathbf{p}}^*$ of the optimal solution can be obtained by solving Problem (15).

(ii) a new element of the set $\mathcal{Q}$ is explored: Assume that an optimal solution of Problem (16) exists and the inequality

$$\hat{z} \;>\; (\mathbf{b} - \mathbf{F}(\hat{\mathbf{r}}))^T \hat{\mathbf{u}^*} + f(\hat{\mathbf{r}}) \tag{21}$$

holds, where $\hat{\mathbf{u}}^*$ the optimal solution of Problem (16). Then, Inequality (18) does not hold for the vector $\left(1, \hat{\mathbf{u}}^{*^T}\right)^T$. Therefore a new candidate for being $(\hat{\mathbf{r}}, \hat{z})$ must be generated by taking into account this new inequality.

(iii) two new elements of the set $\mathcal{Q}$ is explored: Assume that no optimal solution of Problem (16) exists but the objective function is unbounded. Assume that Problem (16) is solved by the simplex method. At the very moment when the unboundedness of the problem is recognized there are a current basic solution and a extremal direction of the unboundedness, say $\hat{\mathbf{u}}$ and $\hat{\mathbf{t}}$, respectively. Then Inequality (18) must be satisfied for the vectors $(1, \hat{\mathbf{u}})$, and $(0, \hat{\mathbf{t}})$, too.

A candidate $(\hat{\mathbf{r}}, \hat{z})$ can be generated as follows. Let $\hat{\mathcal{Q}}$ be the subset of $\mathcal{Q}$ consisting of the explored elements. Then the new candidate is an optimal solution of the problem

$$\begin{array}{c} \max z \\ \forall\,(u_0, \mathbf{u}) \in \hat{\mathcal{Q}}: \; u_0(f(\mathbf{r}) - z) + \mathbf{u}^T(\mathbf{b} - \mathbf{F}(\mathbf{r})) \;\geq\; 0 \\ \mathbf{r} \in \mathcal{S}. \end{array} \tag{22}$$

Thus the Benders decomposition solves the difficult Problem (14) by a finite alternating sequence of Problems (16), and (22), which are of type linear programming and a pure not linear, e.g. in our particular case integer programming.

It is worth to note that as there is no restriction on the set $\mathcal{S}$, if there are any constraints containing only the variables $\mathbf{r}$, then the satisfaction of these constraints can be included in the definition of the set $\mathcal{S}$.

5 The frame of the Benders decomposition in the particular case

This section consists of two parts. First the special structures of the coefficient matrices of problem (13) are explored. On the based of it the particular form of the Benders decomposition is described. Further special properties are discussed in the next section.

In what follows, $\mathbf{e}_t$ is again the t-dimensional vector of all the components of which are 1.

Assume that the order of the components in vector $\mathbf{w}$ is $w_{111}, w_{112}, ..., w_{11n}$, $w_{121}, ..., w_{nnn}$. Similarly let $d_{11}, d_{12}, ..., d_{1n},\ d_{21}, d_{22}, ..., d_{nn}$ be the order of the components in vector $\mathbf{d}$ formed from the distances. It is also assumed that the order of the components in $\mathbf{x}$ is $x_{11}, x_{12}, ..., x_{1n}, x_{21}, x_{22}, ..., x_{nn}$. Then vector $\mathbf{f}$, which is the vector of the objective function coefficients of $\mathbf{w}$, is obtained by the following matrix multiplication:

$$\mathbf{f} = \begin{pmatrix} -\mathbf{e}_n & \mathbf{0} \; ... & \mathbf{0} \\ \mathbf{0} & -\mathbf{e}_n \; ... & \mathbf{0} \\ & ... & \\ \mathbf{0} & \mathbf{0} \; ... & -\mathbf{e}_n \end{pmatrix} \mathbf{d}. \tag{23}$$

The structure of the coefficient matrices of the constraints are as follows. $\mathbf{A}_1$ is the matrix of an $n \times n$ assignment problem, i.e. its structure is this:

$$\mathbf{A}_1 = \begin{pmatrix} \mathbf{A}_{11} \\ \mathbf{A}_{12} \end{pmatrix}, \mathbf{A}_{11} = \begin{pmatrix} \mathbf{e}_n^T & \mathbf{0}^T & ... & \mathbf{0}^T \\ \mathbf{0}^T & \mathbf{e}_n^T & ... & \mathbf{0}^T \\ & & ... & \\ \mathbf{0}^T & \mathbf{0}^T & ... & \mathbf{e}_n^T \end{pmatrix}, \mathbf{A}_{12} = \begin{pmatrix} \mathbf{I}_n & \mathbf{I}_n & ... & \mathbf{I}_n \end{pmatrix}, \tag{24}$$

where $\mathbf{I}_n$ is the $n \times n$ unit matrix. Constraints (3)-(5) describe the feasible set of a TSP. Therefore $\mathbf{B}_2$ consists of three parts, i.e.

$$\mathbf{B}_2 = \begin{pmatrix} \mathbf{B}_{21} \\ \mathbf{B}_{22} \\ \mathbf{B}_{23} \end{pmatrix}, \tag{25}$$

and $\begin{pmatrix} \mathbf{B}_{21} \\ \mathbf{B}_{22} \end{pmatrix}$ is again the matrix of an $n \times n$ assignment problem, i.e.

$$\mathbf{B}_{21} = \mathbf{A}_{11}, \mathbf{B}_{22} = \mathbf{A}_{12}. \tag{26}$$

Constraint (5) excludes short circuits. In principle it excludes all of them, but in practice only the explored ones. Therefore its rows are the negative characteristic vectors of sets of components containing at least 2 and at most $n-1$ components. The following notation is used:

$$\mathbf{B}_{23} = \begin{pmatrix} -\mathbf{v}_1^T \\ ... \\ -\mathbf{v}_m^T \end{pmatrix}, \tag{27}$$

where

$$\mathbf{v}_i \in \{0,1\}^n, \quad 2 \leq \sum_{k=1}^{n} v_{ik} \leq n-1, \quad i = 1, ..., m.$$

The appropriate right-hand side vector, i.e. $\mathbf{b}_2$, is partitioned accordingly, i.e.

$$\mathbf{b}_2 = \begin{pmatrix} \mathbf{b}_{21} \\ \mathbf{b}_{22} \\ \mathbf{b}_{23} \end{pmatrix} = \begin{pmatrix} \mathbf{e}_n \\ \mathbf{e}_n \\ -\mathbf{e}_m \end{pmatrix}. \tag{28}$$

The third set of constraints, i.e. Inequalities (10), and (9), describe the linearization of the $x_{il}y_{kl}$ products. Notice that (9) must be multiplied by -1 to obtain the form used in Problem (13). All the matrices $\mathbf{A}_3$, $\mathbf{B}_3$, and $\mathbf{C}_3$ and the vector $\mathbf{b}_3$ are partitioned according to the two sets of constraints, i.e.

$$\mathbf{A}_3 = \begin{pmatrix} \mathbf{A}_{31} \\ \mathbf{A}_{32} \end{pmatrix}, \quad \mathbf{B}_3 = \begin{pmatrix} \mathbf{B}_{31} \\ \mathbf{B}_{32} \end{pmatrix}, \quad \mathbf{C}_3 = \begin{pmatrix} \mathbf{C}_{31} \\ \mathbf{C}_{32} \end{pmatrix}, \quad \text{and} \quad \mathbf{b}_3 = \begin{pmatrix} \mathbf{b}_{31} \\ \mathbf{b}_{32} \end{pmatrix},$$

where the sizes of $\mathbf{A}_{31}$, and $\mathbf{B}_{31}$ are $n^3 \times n^2$, the size of $\mathbf{C}_{31}$ is $n^3 \times n^3$ and the structure of these matrices is as follows:

$$\mathbf{A}_{31} = \begin{pmatrix} \mathbf{I}_n & \mathbf{0} & \mathbf{0} \ldots & \mathbf{0} \\ & & \ldots & \\ \mathbf{I}_n & \mathbf{0} & \mathbf{0} \ldots & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_n & \mathbf{0} \ldots & \mathbf{0} \\ & & \ldots & \\ \mathbf{0} & \mathbf{I}_n & \mathbf{0} \ldots & \mathbf{0} \\ & & \ldots & \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \ldots & \mathbf{I}_n \\ & & \ldots & \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \ldots & \mathbf{I}_n \end{pmatrix}, \tag{29}$$

$$\mathbf{B}_{31} = \begin{pmatrix} \mathbf{I}_{n^2} \\ \ldots \\ \mathbf{I}_{n^2} \end{pmatrix}, \tag{30}$$

$$\mathbf{C}_{31} = -\mathbf{I}_{n^3}, \tag{31}$$

$$\mathbf{b}_{31} = \mathbf{e}_{n^3}. \tag{32}$$

Furthermore

$$F\mathbf{A}_{32} = -\mathbf{A}_{31}, \quad \mathbf{B}_{32} = -\mathbf{B}_{31}, \quad \mathbf{C}_{32} = -2\mathbf{C}_{31}, \quad \text{and} \quad \mathbf{b}_{32} = \mathbf{0}. \tag{33}$$

The Benders decomposition is applied with the following "casting" of the variables. The role of the linear continuous variables, i.e. the role of the variables $\mathbf{p}$, is assigned to the vector $\mathbf{x}$ and the pair $(\mathbf{y}, \mathbf{w})$ plays the role of the vector $\mathbf{r}$.

As it has great importance, the set $\mathcal{S}$ is given in a separated definition.

Definition 51 *The set $\mathcal{S}$ is defined such that both $\mathbf{y}$, and $\mathbf{w}$ are binary vectors satisfying:*

- $w_{ikl} = 1$ *only if* $y_{kl} = 1$ *and*
- *the vector* $\mathbf{y}$ *describes a Hamiltonian circuit.*

During the algorithm the requirement that $\mathbf{y}$ must define a Hamiltonian circuit is handled dynamically, i.e. only the constraints are required that exclude a potential non-Hamiltonian solution.

It is supposed that the vector $\mathbf{u}$ of the variables of Problem (16) is partitioned into seven parts of the constraints. The seven parts are the two parts of the assignment problem, i.e. (1) and (2), the assignment part of the TSP and the exclusion of the small circuits, i.e. (3) and (4) together and (5), finally (9) and (10). Then the particular form of Problem (16) for a fixed pair $(\hat{\mathbf{y}}, \hat{\mathbf{w}})$ is:

$$\begin{aligned}
\min & \left(\begin{pmatrix} \mathbf{b}_1 \\ \mathbf{b}_2 \\ \mathbf{b}_3 \end{pmatrix} - \begin{pmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{B}_2 & \mathbf{0} \\ \mathbf{B}_3 & \mathbf{C}_3 \end{pmatrix} \begin{pmatrix} \hat{\mathbf{y}} \\ \hat{\mathbf{w}} \end{pmatrix} \right)^T \mathbf{u} \\
\text{subject to} & \\
& (\mathbf{A}_1^T, \mathbf{0}, \mathbf{A}_3^T)\mathbf{u} \geq \mathbf{d} \\
& \mathbf{u}_{23} \geq \mathbf{0} \\
& \mathbf{u}_3 \geq \mathbf{0}.
\end{aligned} \tag{34}$$

The particular form of the individual linear inequalities is

$$u_{11i} + u_{12l} + \sum_{k=1}^{n} u_{31ikl} - \sum_{k=1}^{n} u_{32ikl} \geq -d_{il}. \tag{35}$$

Hence $\mathbf{u} = \mathbf{0}$ is always a feasible solution of (34), as the distances are non-negatives. It means that the assumption that (16) has a feasible solution is automatically satisfied.

In the description of the algorithm the following notations are used:

- $\mathcal{R}$ is the set of pairs of binary vectors $(\mathbf{y}, \mathbf{w})$ satisfying that $y_{kl} = 0$ implies that $\forall\, i : w_{ikl} = 0$.
- If (T) denotes an optimization problem then let $\text{OPT}(T)$ denote an optimal solution of (T) provided by any algorithm used to solve the problem.
- Similarly if (T) is a linear programming problem then $\text{EXTR}(T)$ is the last extremal point visited by the simplex method and
- $\text{DIRECTION}(T)$ is the direction such that the value of the objective function is unbounded on the half-line started from $\text{EXTR}(T)$ along the direction $\text{DIRECTION}(T)$.
- The set $\mathcal{H}$ consists of smaller circuits that appeared in a vector $\mathbf{y}$, i.e. the appropriate constraints (5) must be required for each $H \in \mathcal{H}$.
- For a given vector $\mathbf{y}$ $\text{SUBCIRCUIT}(\mathbf{y})$ is one small, i.e. non-Hamiltonian circuit appearing in $\mathbf{y}$ and
- $\text{CIRCUIT}(\mathbf{y})$ is the number of circuits represented by $\mathbf{y}$. It is 1 if and only if $\mathbf{y}$ represents a Hamiltonian circuit.
- The variables depending on the iteration are the following:
  - β the index of the iteration,

- $\mathcal{Q}_\beta$ the set of explored extremal points and directions,
- z_β the optimal objective function value of linear programming subproblem,
- s_β the optimal value of the integer programming subproblem,
- $(\mathbf{y}_\beta, \mathbf{w}_\beta)$ the optimal solution of the integer programming subproblem denoted by INTEGER$_\beta$,
- $(\mathbf{x}^*, \mathbf{y}^*, \mathbf{w}^*)$ the optimal solution of the original problem,
- $\mathbf{u}_\beta$ the extremal point obtained in the k-th iteration,
- $\mathbf{v}_\beta$ the extremal direction obtained in the k-th iteration.

Algorithm 5.1

1. **Begin**
2. $\mathcal{Q}_0 := \emptyset$
3. $\mathcal{H} := \emptyset$
4. $s_0 := +\infty$
5. $z_0 := 0$
6. $(\mathbf{y}_0, \mathbf{w}_0) \in \mathcal{S}$ {An arbitrary element}
7. $\beta := 0$
8. $$z_\beta = \min \left(\begin{pmatrix} \mathbf{b}_1 \\ \mathbf{b}_2 \\ \mathbf{b}_3 \end{pmatrix} - \begin{pmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{B}_2 & \mathbf{0} \\ \mathbf{B}_3 & \mathbf{C}_3 \end{pmatrix} \begin{pmatrix} \mathbf{y}_\beta \\ \mathbf{w}_\beta \end{pmatrix} \right)^T \mathbf{u}$$
   SUBJECT TO $\quad (34_\beta)$
   $$(\mathbf{A}_1^T, \mathbf{0}, \mathbf{A}_3^T)\mathbf{u} \geq \mathbf{d} \quad \mathbf{u}_{23} \geq \mathbf{0} \quad \mathbf{u}_3 \geq \mathbf{0}$$
9. **if** $z_\beta \geq s_\beta - \mathbf{f}^T \mathbf{w}_\beta$
10. **then**
11. **begin**
12. **goto 36**.
13. **end**
14. **if** $-\infty < z_\beta < s_\beta - \mathbf{f}^T \mathbf{w}_\beta$
15. **then**
16. **begin**
17. $\mathbf{u}_\beta :=$OPT(34_β)
18. $\mathcal{Q}_{k+1} := \mathcal{Q}_\beta \cup \left\{ \begin{pmatrix} 1 \\ \mathbf{u}_\beta \end{pmatrix} \right\}$
19. **end**
20. **else**
21. **if** $z_\beta = -\infty$
22. **then**
23. **begin**
24. $\mathbf{u}_\beta :=$EXTR(34_β)
25. $\mathbf{v}_\beta :=$DIRECTION(34_β)
26. $\mathcal{Q}_{\beta+1} := \mathcal{Q}_\beta \cup \left\{ \begin{pmatrix} 1 \\ \mathbf{u}_\beta \end{pmatrix}, \begin{pmatrix} 0 \\ \mathbf{v}_\beta \end{pmatrix} \right\}$
27. **end**

28. **repeat**
29. $s_{\beta+1} := \max s$
SUBJECT TO
$\forall (u_0, \mathbf{u}) \in \mathcal{Q}_\beta$:

$$u_0(\mathbf{f}^T\mathbf{w} - s) + (\mathbf{u}_1^T, \mathbf{u}_2^T, \mathbf{u}_3^T) \left(\begin{pmatrix} \mathbf{e}_{2n} \\ \mathbf{b}_2 \\ \mathbf{b}_3 \end{pmatrix} - \begin{pmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{B}_2 & \mathbf{0} \\ \mathbf{B}_3 & \mathbf{C}_3 \end{pmatrix} \begin{pmatrix} \mathbf{y} \\ \mathbf{w} \end{pmatrix} \right) \geq 0$$

(INTEGER$_\beta$)

$$(\mathbf{y}, \mathbf{w}) \in \mathcal{R}$$
$$\forall H \in \mathcal{H}: \sum_{k \in H} \sum_{l \in \overline{H}} y_{kl} \geq 1$$

30. $(\mathbf{y}_{\beta+1}, \mathbf{w}_{\beta+1})$:=OPT(INTEGER$_\beta$)
31. **if** CIRCUIT$(\mathbf{y}_{\beta+1}) > 1$
32. **then** $\mathcal{H} := \mathcal{H} \cup \{$SUBCIRCUIT$(\mathbf{y}_\beta)\}$
33. **until** CIRCUIT$(\mathbf{y}_{\beta+1}) > 1$
34. $\beta := \beta + 1$
35. **goto 8**.
36. $\mathbf{y}^* := \mathbf{y}_\beta$
37. $\mathbf{w}^* := \mathbf{w}_\beta$
38. $\mathbf{x}^* :=$ OPT$(\max \{(-\mathbf{d})^T\mathbf{x} \mid \mathbf{A}_1\mathbf{x} = \mathbf{e}_{2n},\ \mathbf{A}_3\mathbf{x} \leq \mathbf{b}_3 - \mathbf{B}_3\mathbf{y}^* - \mathbf{C}_3\mathbf{w}^*,\ \mathbf{x} \in \{0,1\}^{n \times n}\})$.
39. **end**

The correctness of the algorithm has not yet been proved as the problem in Row 38 giving the optimal $\mathbf{x}$ part of the solution is a combinatorial optimization problem instead of a pure linear programming one. The aim of the next section is to prove that the current version of (15) reserves its combinatorial nature.

6 The combinatorial nature of the Benders decomposition in the particular case

The particular form of (15) is

$$\begin{aligned} \max\ & (-\mathbf{d})^T\mathbf{x} \\ & \mathbf{A}_1\mathbf{x} = \mathbf{e}_{2n} \\ & \mathbf{A}_3\mathbf{x} \leq \mathbf{b}_3 - \mathbf{B}_3\hat{\mathbf{y}} - \mathbf{C}_3\hat{\mathbf{w}} \\ & \mathbf{x} \in \{0,1\}^{n^2}. \end{aligned} \tag{36}$$

Without the inequalities in the third row Problem (36) is an assignment problem. The objective of this section is just to show that (36) behaves in the frame of the Benders decomposition like an assignment problem, i.e. although it is defined as an integer programming problem, it can be handled as a linear programming problem. For the possible values of each pair of (y_{kl}, w_{ikl}) there are the following cases considering the appropriate constraints (9), (10) and the fact that $(\mathbf{y}, \mathbf{w}) \in \mathcal{R}$.

y_{kl}	w_{ikl}	x_{il}	the binding constraint
1	1	1	(10)
1	0	0	(9)
0	1	–	the case cannot occur
0	0	0,1	no binding constraint

These constraints may cause two types of infeasibilities. Then the appropriate sample of Problem (16) is unbounded. It is shown below that in both cases it is possible to give a direction such that the objective function of (16) is unbounded along it. To do so the following form of Farkas lemma is used.

Lemma 61 *Let* $\mathbf{G}$ *be an* $m \times n$ *matrix and* $\mathbf{g}$ *an* m*-dimensional vector. The* n*-dimensional vector of variables is denoted by* $\mathbf{t}$*. If the system*

$$\mathbf{Gt} \;\leq\; \mathbf{g} \tag{37}$$

has no solution then there is a nonnegative m*-dimensional vector* $\underline{\lambda}$ *such that*

$$\mathbf{G}^T\underline{\lambda} \;=\; \mathbf{0} \quad \textit{and} \quad \mathbf{g}^T\underline{\lambda} \;<\; 0. \tag{38}$$

Assume that the linear programming problem

$$\min \mathbf{g}^T\underline{\mu}$$
$$\mathbf{G}^T\underline{\mu} = \mathbf{d}$$

is to be solved, where $\mathbf{d}$ is a fixed n-dimensional vector. Then if system (38) has a solution then the linear programming problem either has no feasible solution, or is unbounded. In the latter case the vector $\underline{\lambda}$ gives a direction such that starting from any feasible solution the objective function is unbounded along this direction.

When lemma 61 is applied to (36) then the particular form of the system (37) is:

$$\mathbf{A}_1\mathbf{x} = \mathbf{e}_{2n}, \quad \mathbf{A}_3\mathbf{x} \leq \mathbf{b}_3 - \mathbf{B}_3\hat{\mathbf{y}} - \mathbf{C}_3\hat{\mathbf{w}}, \quad -\mathbf{x} \leq \mathbf{0}, \tag{39}$$

i.e. in this particular case the matrix $\mathbf{G}$ is

$$\mathbf{G} \;=\; \begin{pmatrix} \mathbf{A}_1 \\ \mathbf{A}_3 \\ -\mathbf{I}_{n\times n} \end{pmatrix}. \tag{40}$$

Similarly

$$\mathbf{g} \;=\; \begin{pmatrix} \mathbf{e}_{2n} \\ \mathbf{b}_3 - \mathbf{B}_3\hat{\mathbf{y}} - \mathbf{C}_3\hat{\mathbf{w}} \\ \mathbf{0} \end{pmatrix}. \tag{41}$$

As the first set of constraints is an equation system, it is allowed that their multipliers take negative values, too.

Case 1. Too many 1's are required. Assume that there are indices i_1, i_2, k_1, k_2, l with $i_1 \neq i_2$ such that $\hat{y}_{k_1l} = \hat{y}_{k_2l} = \hat{w}_{i_1k_1l} = \hat{w}_{i_2k_2l} = 1$ then the sum

$$\sum_{i=1}^{n} x_{il}$$

is at least 2 contradicting to the corresponding constraint (2). Then the appropriate inequalities of type (9) are

$$-x_{i_1k_1} - \hat{y}_{k_1l} + 2\hat{w}_{i_1k_1l} \leq 0, \quad -x_{i_2k_2} - \hat{y}_{k_2l} + 2\hat{w}_{i_2k_2l} \leq 0,$$

which are equivalent to

$$-x_{i_1k_1} \leq -1, \quad -x_{i_2k_2} \leq -1 \tag{42}$$

according to the current value of $\hat{\mathbf{y}}$, and $\hat{\mathbf{w}}$. Then the non-zero components of the appropriate $\underline{\lambda}$ vector are as follows. The weight of the equation is

$$\sum_{i=1}^{n} x_{il} \;=\; 1,$$

and of the two inequalities of (42), and finally of the nonnegativity constraints

$$-x_{il} \;\leq\; 0 \quad i = 1, 2, ..., n, \;\; i \neq i_1, i_2$$

are 1. The weight of all other constraints is 0. With this weight Relation (38) is obtained. At the same time a direction of unboundedness of Problem (16) is determined, which (with a zero first component) must be added to the set $\mathcal{Q}$. It is easy to check if this case occurs. If the answer is yes then command in Row 25 of the algorithm can be executed without applying any linear programming solver. The command in row 24 can be temporarily omitted as the extremal point can be added later to the set $\mathcal{Q}$ (with the supplementary component 1). The explanation is this. The scheme of Benders decomposition does not determine the order in which the constraints of type (18) must be claimed. The only point is that in each iteration at least one new constraint must be added to Problem (22).

If there are indices i, k_1, k_2, l_1, l_2 such that $l_1 \neq l_2$ and $\hat{y}_{k_1l_1} = \hat{y}_{k_2l_2} = \hat{w}_{ik_1l_1} = \hat{w}_{ik_2l_2} = 1$ one can get Relation (38) in a similar way. It worth to note that this type of infeasibility does not exist with $k = k_1 = k_2$ because then the vector $\mathbf{y}$ is not a characteristic vector of a Hamiltonian circuit.

Case 2. Too few 1's are allowed. Here it is supposed that Case 1 does not occur. Not all x_{il} might be 1 as in the case $y_{kl} = 1$, $w_{ikl} = 0$ x_{il} must be 0. Let $\mathcal{P}$

be the set of such pairs. The elements of $\mathcal{P}$ are called *prohibited pairs of indices.* In any feasible solution of the original problem the matrix $\mathbf{x}$ must be such that it contains in each row and in each column exactly one 1 and all other elements are 0. This requirement can be satisfied only if the maximal solution of the following matching problem consists of n edges. Let $\mathcal{V} = \{1, 2, ..., n\} \cup \{\hat{1}, \hat{2}, ..., \hat{n}\}$ be the set of vertices. The set of edges is $\mathcal{E} = \{(i, \hat{j}) \mid 1 \leq i, j \leq n\} \setminus \mathcal{P}$. Kœnig's theorem says that a matching of n edges exist if and only if for every nonempty subset $\mathcal{S}$ of $\{1, 2, ..., n\}$ the relation

$$|\mathcal{S}| \;\leq\; |\{\hat{j} \mid \exists i \in \mathcal{S} : \, (i, \hat{j}) \in \mathcal{E}\}|$$

holds.

The matching problem can be solved by a polynomial algorithm. If the optimal value is n then Problem (15) has an optimal solution. If Case 1 does not occur then still some variables x_{il} might be fixed to 1 but no other variable is fixed to 1 in its row, and column. These fixings must be taken into consideration when the matching problem is solved.

If the optimal value of the matching problem is less than n then the multipliers in (38) are these. Then there is a nonempty index set $\mathcal{S} \subset \{1, 2, ..., n\}$ and another set $\mathcal{T} \subset \{\hat{1}, \hat{2}, ..., \hat{n}\}$ such that $|\mathcal{S}| > |\mathcal{T}|$ and $\forall\, i \in \mathcal{S} \,\forall\, \hat{j} \in \{\hat{1}, \hat{2}, ..., \hat{n}\} \setminus \mathcal{T}$: $(i, \hat{j}) \notin \mathcal{E}$. The multipliers of Equations (1) belonging to indices $i \notin \mathcal{S}$ are 1. The multipliers of Equations (2) belonging to an index $\hat{j} \notin \mathcal{T}$ are -1. The current form of Constraints (9) for prohibited pair, i.e. if $y_{kl} = 1$ and $w_{ikl} = 0$, is $x_{il} \leq 0$. As all pair with $(i, \hat{j})$ with $i \in \mathcal{S}$ and $\hat{j} \notin \mathcal{T}$ are prohibited, therefore the multipliers of all such constraints of type (9) are 1. Finally the multipliers of the nonnegativity constraints of variables $x_{i\hat{j}}$ with $i \notin \mathcal{S}$ and $\hat{j} \in \mathcal{T}$ are 1. The multipliers of all other constraints are 0.

Thus Problem (15) or equivalently (34_β) can be handled during the algorithm as follows:

- At first Case 1 type of infeasibilities are eliminated.

-Then Problem (36) is reduced according to which x_{il}'s must be 1. The reduced problem is solved with the following modified objective function $(-\hat{\mathbf{d}}^T)\mathbf{x}$, where

$$-\hat{d}_{il} \;=\; \begin{cases} -d_{il} \text{ if } (i, l) \notin \mathcal{P} \\ -\infty \text{ if } (i, l) \in \mathcal{P}. \end{cases}$$

Thus an assignment problem is obtained and it can be solved by some combinatorial algorithm. If the optimal value is finite then it is the optimal value of the current Problem (36). Optimal solution can be generated e.g. via complementary slackness. If the optimal value is $-\infty$ then Case 2 type of infeasibility occurs and a direction of unboundedness is obtained.

References

[1] Benders, J. F., Partitioning Procedures for Solving Mixed-Variables Programming Problems, *Numerische Mathematik*, 4 (1962), 238–252.

[2] Yves Crama, J. van de Klundert, F.C.R. Spieksma, Production planning problems in printed circuit board assembly, *Discrete Applied Mathematics*, 123 (2002), 339–361.

[3] König Dénes, Graphok és mátrixok, *Mat. és Fiz. Lapok 38. 1931.*

Invariant cones and polyhedra for dynamical systems

Zoltán HORVÁTH[⋆] [⋆⋆]

Department of Mathematics
Széchenyi István University,
Egyetem tér 1, 9026 Győr, Hungary

Abstract. In this paper we study the mathematical modeling of some unsteady physical processes by time dependent differential equations and their numerical approximations. The focus is on examining whether some physical properties, such as invariance in time of some convex, closed polyhedral sets of the state space by the state variables were preserved in the mathematical modeling, particularly at the numerical solution of the corresponding time dependent differential equations. We determine time step sizes for Runge-Kutta time discretization methods by a practically simple and useful formula that guarantee the discrete invariance property. From the results we can see that the existence of such positive step sizes is not automatically fulfilled for high order and stable methods.

AMS 2000 subject classifications: 37L25, 37N30, 65L06, 35B99

Key words and phrases: discrete and continuous dynamical system; positively invariant set; numerical methods; Runge-Kutta methods

1 Motivation and scope of the paper

Mathematical modeling of physical processes often leads to differential equations (DEs). For a proper modeling we need not only a good approximation of the physical quantities by the mathematical variables, which are typically solutions or some functions of the solutions of the DE, but we want to preserve the basic characteristics of the quantities of the physical process.

As an example, consider the unsteady heat conduction in a certain inhomogeneous body with known thermal diffusivity function $\sigma > 0$ of space and suppose that there are no sinks nor sources of heat, moreover the temperature is maintained fixed T_b at the boundary of the body.

We know from the basic principles of thermodynamics that in such situations heat can not flow from a colder place to a warmer one. Hence in all time level the temperature of the coldest point and consequently the minimum of the temperature increases (or at least not decreases) with time. Similarly, the maximal temperature of the body decreases (not increases) forward in time. More formally, let $T(t, x)$ denote the temperature at time t and space point x of the body Ω, which is, for simplicity, supposed to be one dimensional of length ω;

[⋆] email: horvathz@sze.hu

[⋆⋆] Supported by OTKA (Hungarian Scientific Research Fund) T043177.

then we have

$$\forall t > 0 : \min_{x \in \Omega} T(t,x) \geq \min_{x \in \Omega} T(0,x) =: T_{0,min}, \tag{1}$$

$$\forall t > 0 : \max_{x \in \Omega} T(t,x) \leq \max_{x \in \Omega} T(0,x) =: T_{0,max}. \tag{2}$$

The basic mathematical model of this heat conduction process consists of a parabolic partial differential equation (PDE) for the function $u(t,x)$ representing the scaled temperature $T(t,x) - T_b$ in the mathematical model, subject to boundary and initial conditions: $\frac{\partial u}{\partial t} = \frac{\partial}{\partial x}(\sigma(x)\frac{\partial u}{\partial x})$; for $t > 0, x \in [0,\omega]$ the boundary condition reads $u(t,0) = u(t,\omega) = 0$ ($\forall t \geq 0$) and the initial temperature $u(0,.)$ is given.

We know (see e.g. [12], Example VII.5.2.) that this model preserves the examined thermodynamical properties corresponding to (1) and (2), namely

$$\forall t > 0 : \min_{x \in [0,\omega]} u(t,x) \geq \min_{x \in [0,\omega]} u(0,x) =: u_{0,min}, \tag{3}$$

$$\forall t > 0 : \max_{x \in [0,\omega]} u(t,x) \leq \max_{x \in [0,\omega]} u(0,x) =: u_{0,max}. \tag{4}$$

The mathematical model based on the PDE is useful to answer several questions of physics. However, for quantitative results we usually has to approximate u numerically. This means that we prescribe discrete time points $t_0 = 0 < t_1 < ... < t_M$ and points in space $x_1, x_2, ..., x_N \in \Omega$ and approximate the solution of the PDE problem at these points as $u(t_n, x_k) \approx u_{n,k}$ for all n, k.

One way of deriving such a full discretization may be obtained in two steps by the method of lines: first, in the so-called semi-discretization we discretize the space variable only as $u(t, x_k) \approx U_k(t)$ ($t \geq 0, k = 1, ..., N$) and then, in the second step, $U_k(t)$ is approximated further at discrete time points $t = t_n$ by $u_{n,k}$. This finally gives $u(t_n, x_k) \approx U_k(t_n) \approx u_{n,k}$.

In the semi-discretization we can approximate, for example, the partial derivatives in the space variable by some difference formulas based on the $u(t, x_k) \approx U_k(t)$ values. This issues in a system of ordinary differential equations (ODEs) subject to an initial condition for the vector valued function U with $U(t) = (U_1(t), ..., U_N(t))^T$ (for all $t \geq 0$); this problem can be considered the mathematical model of the physical process as well. One of these ODE systems may be obtained by using two-point central differences resulting in the ODE $U'(t) = LU(t)$ with a tridiagonal matrix L with non-negative off-diagonal entries and dominant diagonal (see [3], [10]). For the validity of the examined physical principle we need that $U(0) \in \mathcal{C} := [u_{0,min}, u_{0,max}]^N$ should imply $U(t) \in \mathcal{C}$ for all $t > 0$; and this can be proven in fact.

In the next step we approximate the solution of the ODE subject to the initial condition numerically, for example by a Runge-Kutta method (for definition see Section 4.1) below. This produces approximate values $u_n = (u_{n,1}, ..., u_{n,N})^T$

to $U(t_n)$ for all n. To ensure the physical principle be in force we need that $u_0 \in \mathcal{C}$ should imply $u_n \in \mathcal{C}$ for all $n = 1, 2, ...$ ("discrete forward invariance").

The goal of the paper is to examine this discrete invariance property in a much more general settings: we consider system of non-linear differential equations, discrete invariance of $\mathcal{C}$ when it is a convex, closed polyhedral set and arbitrary Runge-Kutta methods for discretization. For such problems we are to find and analyse conditions on the time step sizes of the method that guarantee the discrete invariance of $\mathcal{C}$.

2 Definitions, notations

We denote the state space of the dynamical system by V. For simplicity we assume that $V = \mathbb{R}^N$, although most of the results of the paper hold true for general Banach spaces in the present form. Let $f : V \to V$ be a fixed continuously differentiable function. The semiflow $\Phi : [0, \infty) \times V \to V$ is generated by the differential equation with right hand side function f

$$u'(t) = f(u(t)), \quad t \geq 0. \tag{5}$$

For simplicity we suppose that (5) equipped with the initial condition $u(0) = u_0$ has a unique solution $u : [0, \infty) \to V$ for all $u_0 \in V$. Then, with this solution, $\Phi(t, u_0) := u(t)$ $(t \in [0, \infty), u_0 \in V)$.

We call the subset $\mathcal{C}$ of V positively (or forward) invariant w.r.t. Φ iff the trajectories emanating from $\mathcal{C}$ remain completely in $\mathcal{C}$, namely iff

$$\forall u_0 \in \mathcal{C} \ \forall t \geq 0 \ : \ \Phi(t, u_0) \in \mathcal{C}.$$

Positive invariance of closed, convex sets can be characterized by their tangent cones, see Lemma 2 below. The tangent cone (see also [1] and the references therein) of $\mathcal{C}$ at $c \in V$ is $\mathcal{T}(c) = \mathcal{T}_{\mathcal{C}}(c)$ with

$$\mathcal{T}_{\mathcal{C}}(c) := \{ z \in V \mid \liminf_{h \to 0+} \frac{\mathrm{dist}(c + hz, \mathcal{C})}{h} = 0 \} \tag{6}$$

where $\mathrm{dist}(x, \mathcal{C}) := \inf_{y \in \mathcal{C}} \|x - y\|$. Clearly, $\mathcal{T}_{\mathcal{C}}(c)$ equals V or $\emptyset$ whenever c is an interior or exterior point of V, respectively. Hence $\mathcal{T}_{\mathcal{C}}(c)$ can be non-trivial only if $c \in \partial\mathcal{C}$, i.e. at the boundary points of $\mathcal{C}$. In the latter case $\mathcal{T}_{\mathcal{C}}(c)$ contains the vectors which point to the direction of $\mathcal{C}$ or are tangent to $\mathcal{C}$ when applied at c (for an illustration see Figure 1).

In this paper we shall investigate whether special polyhedral convex sets, e.g. cones and polyhedra, are positively invariant w.r.t. a semiflow Φ. We call a subset of V polyhedral if it is the intersection of finitely many closed halfspaces; the compact polyhedral sets are called polyhedra. Hence, making use of the notation $\mathcal{P}_{P,p} := \{x \in V \mid Px \geq p\}$ for $P \in \mathbb{R}^{K \times N}$ and $p \in \mathbb{R}^K$, $\mathcal{C}$ is a

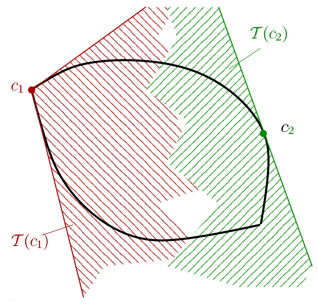


Fig. 1: The tangent cones of a convex set at two points.

- polyhedral set iff $\mathcal{C} = \mathcal{P}_{P,p}$ with some $P \in \mathbb{R}^{K\times N}$ and $p \in \mathbb{R}^K$;
- polyhedral cone iff $\mathcal{C} = \mathcal{P}_{P,0}$ with $P \in \mathbb{R}^{K\times N}$;
- polyhedron iff $\mathcal{C} = \mathcal{P}_{P,p}$ with $P \in \mathbb{R}^{K\times N}$, $p \in \mathbb{R}^K$ whenever $\nexists x_1, x_2 \in \mathcal{C}$: $P(x_1 - x_2) \geq 0$.

Lemma 1. *1. If $\mathcal{C}$ is convex, $c \in \mathcal{C}$, $z \in V$, $\varepsilon > 0$, $c + \varepsilon z \in \mathcal{C}$, then $c + \widetilde{\varepsilon} z \in \mathcal{C}$ holds for all $\widetilde{\varepsilon} \in (0, \varepsilon)$.*
2. Let $\mathcal{C}$ be polyhedral and $c \in \partial\mathcal{C}$. Then $z \in \mathcal{T}(c)$ iff $\exists \varepsilon > 0 :\ c + \varepsilon z \in \mathcal{C}$.

Proof. The first assertion follows from the convexity of $\mathcal{C}$ and the identity $c + \widetilde{\varepsilon} z = \frac{\widetilde{\varepsilon}}{\varepsilon}(c + \varepsilon z) + (1 - \frac{\widetilde{\varepsilon}}{\varepsilon})c$, while the second one from the definitions. □

3 Positively invariant sets for continuous-time semiflows

In this section first we present a necessary and sufficient condition, (7), which means that a convex, closed set $\mathcal{C}$ is positively invariant w.r.t. Φ. Then this condition is simplified equivalently for polyhedra resulting in (8).

Lemma 2. (necessary and sufficient condition for positive invariance)**.** *Let $\mathcal{C}$ be convex and closed. Then $\mathcal{C}$ is positively invariant w.r.t. Φ iff*

$$\forall c \in \partial\mathcal{C} :\ \ f(c) \in \mathcal{T}_{\mathcal{C}}(c). \tag{7}$$

Moreover, if $\mathcal{C}$ is a polyhedron then $\mathcal{C}$ is positively invariant w.r.t. Φ iff

$$\exists \varepsilon > 0 : \quad \forall c \in \mathcal{C} \qquad c + \varepsilon f(c) \in \mathcal{C}. \tag{8}$$

Proof. The first assertion of the lemma is called Nagumo's lemma and its proof can be found e.g. in [4]; see also the references in [1].

Now suppose that $\mathcal{C}$ is a polyhedron. Hence, by Lemma 1, condition (7) is equivalent with

$$\forall c \in \partial\mathcal{C} \quad \exists \varepsilon > 0 : \quad c + \varepsilon f(c) \in \mathcal{C}.$$

Such an $\varepsilon > 0$ exists for $c \in \operatorname{int}\mathcal{C}$ as well. By compactness of $\mathcal{C}$ and continuity of f and hence that of $\varepsilon(c) := \sup\{\varepsilon_1 \in (0,1] | c + \varepsilon_1 f(c) \in \mathcal{C}\}$ condition (8) follows directly. □

3.1 Determination of a suitable $\varepsilon > 0$ for condition (8)

For practical applications it is advantegeous to know the size of ε for which condition (8) holds, see Theorem 1 below. The following lemma constructs such a suitable ε for polyhedral sets with non-obtused face angles. Hence this lemma applies for several important cases such as the non-negative orthant (and the other orthants), rectangles, acute simpleces.

Lemma 3. *Let $\mathcal{C}$ be polyhedral with representation $\mathcal{C} = \mathcal{P}_{P,p}$ $P \in \mathbb{R}^{K\times N}$, $p \in \mathbb{R}^K$, and let $\varphi_i := (P_{i1}, ..., P_{iN})^T$ denote the ith row of P. Suppose that $\mathcal{C}$ is positively invariant w.r.t. Φ and $\varphi_i^T \varphi_j \le 0$ for all $i \neq j$.*

Then condition (8) holds with $\varepsilon = \frac{1}{\ell}$ whenever

$$\forall\, x \in \mathcal{C},\ \forall\, i \in \{1, ..., K\}\ :\ -\ell \le \frac{\varphi_i^T f'(x)\varphi_i}{\varphi_i^T \varphi_i} \le 0. \tag{9}$$

Proof. Let $\varepsilon = 1/\ell$ with ℓ defined in (9) and let $c \in \mathcal{C}$ and $i \in \{1, ..., K\}$ arbitrary and fixed; define

$$\widetilde{c} := c - \frac{\varphi_i^T c - p_i}{\varphi_i^T \varphi_i}\varphi_i.$$

First we show that $\varphi_i^T \widetilde{c} = p_i$ and $\widetilde{c} \in \mathcal{C}$. Indeed, $\varphi_i^T \widetilde{c} = p_i$ is trivial and for any $j \neq i$ there holds $\varphi_j^T \widetilde{c} - p_j = (\varphi_j^T c - p_j) - (\varphi_i^T c - p_i)\varphi_j^T \varphi_i / \varphi_i^T \varphi_i \ge 0$ since the terms in the round brackets are non-negative due to $c \in \mathcal{C}$ and $\varphi_j^T \varphi_i \le 0$ by the conditions of the lemma.

Therefore $\widetilde{c} \in \partial\mathcal{C}$ and thus, by the positive invariance of $\mathcal{C}$ and Lemma 2, there exists $\varepsilon > 0$ with $\widetilde{c} + \varepsilon f(\widetilde{c}) \in \mathcal{C}$; this implies $0 \le \varphi_i^T(\widetilde{c} + \varepsilon f(\widetilde{c})) - p_i = \varepsilon \varphi_i^T f(\widetilde{c})$, i.e. $\varphi_i^T f(\widetilde{c}) \ge 0$.

Then we have, by using the vector $\widetilde{c}$ and applying the mean value theorem to the function $\varphi_i^T f$ and variables c and $\widetilde{c}$ we obtain

$$\begin{aligned}
\varphi_i^T(c + \varepsilon f(c)) - p_i &= (\varphi_i^T c - p_i) + \varepsilon \left(\varphi_i^T f(c) - \varphi_i^T f(\widetilde{c})\right) + \varepsilon \varphi_i^T f(\widetilde{c}) \\
&= (\varphi_i^T c - p_i) + \varepsilon \varphi_i^T f'(x)\varphi_i \frac{\varphi_i^T c - p_i}{\varphi_i^T \varphi_i} + \varepsilon \varphi_i^T f(\widetilde{c}) \\
&= \left(1 + \varepsilon \frac{\varphi_i^T f'(x)\varphi_i}{\varphi_i^T \varphi_i}\right)(\varphi_i^T c - p_i) + \varepsilon \varphi_i^T f(\widetilde{c})
\end{aligned}$$

with some x from the segment joining c and $\widetilde{c}$, which proves $\varphi_i^T(c + \varepsilon f(c)) - p_i \ge 0$. $\square$

Remark 1. One can observe the Rayleigh quotient of $f'(x)$ in (9). Hence the infimum of the smallest eigenvalues of the Jacobian matrices $f'(x), x \in \mathcal{C}$ is suitable for the first inequality of (9) and the infimum is certainly finite for compact polyhedral sets. However, usually one can find a smaller ℓ for (9).

4 Discretization of Φ

Suppose that we are given a time grid $0 = t_0 < t_1 < ...$ with time step sizes $h_n := t_{n+1} - t_n$. Let us approximate the trajectories of the dynamical system at time points t_n as

$$u_n \approx \Phi(t_n, u_0), \qquad n = 0, 1, 2, ..., \; u_0 \in V.$$

We consider "one-step" discretizations, i.e. approximations when

$$u_{n+1} = \Psi(h_n, u_n), \quad n \geq 0, \qquad u_0 \text{ given.} \tag{10}$$

Here $\Psi : [0, H_{\text{def}}] \times V \to V$ depends on the discretization method and f. We assume that Ψ is well-defined over $[0, H_{\text{def}}] \times V$.

Definition 1. $\mathcal{C}$ is discrete positively invariant under the one-step method with definition (10) (or shortly: under Ψ) with step size constant $H \in (0, H_{\text{def}}]$ iff

$$\forall u_0 \in \mathcal{C} \,:\, (\forall n : \, h_n \in [0, H] \; \Rightarrow \; \forall n : \, u_n \in \mathcal{C})$$

whenever the sequence u_n is defined by (10).

The latter is equivalent with the condition that $\Psi(h, .)$ maps $\mathcal{C}$ into itself for all $h \in [0, H]$:

$$\forall u \in \mathcal{C}, \; \forall h \in [0, H] \,:\, \Psi(h, u) \in \mathcal{C}.$$

4.1 Time discretization of Φ with Runge-Kutta methods

Runge-Kutta (RK) methods are well-known and widely used one-step methods for discretization of semiflows, see e.g. [3].

The RK method with coefficient arrays $A = (a_{ij}) \in \mathbb{R}^{s \times s}$ and $b = (b_i) \in \mathbb{R}^s$ is given by (10) with

$$\Psi(h, u) = u + h \sum_{i=1}^{s} b_i \, f(y_i)$$

where $y_i, \; i = 1, \ldots, s$ form the (unique) solution of the system of algebraic equations

$$y_i = u + h \sum_{j=1}^{s} a_{ij} \, f(y_j) \quad i = 1, ..., s. \tag{11}$$

This method being determined completely by its arrays is denoted by $\text{RK}(A, b)$. During this paper we assume that system (11) possesses a unique solution whenever $u \in V$ and $h \in [0, H_{\text{def}}]$ are arbitrary. For more details about unique solvability and formulas for H_{def} in terms of some other characteristics, esp. measures of dissipativity of f see e.g. [3].

We remark that A and b are usually chosen according to the stability and accuracy of the method.

We shall see below in the next sections that the scheme functions and the absolute monotonicity radius of $\text{RK}(A, b)$ play an important role in our analysis.

Definition 2. The scheme functions of RK(A,b) are defined as

$$\begin{aligned} K_A(z) &:= (I - zA)^{-1}e, & J_A(z) &:= A(I - zA)^{-1}, \\ K_b(z) &:= 1 + b^T z(I - zA)^{-1}e, & J_b(z) &:= b^T (I - zA)^{-1} \end{aligned}$$

where z is a real variable, $e = (1, ..., 1)^T \in \mathbb{R}^s$ and I denotes here and throughout the paper the identity matrix of appropriate order (here it is of $s \times s$ order).

Further, the absolute monotonicity radius of RK(A,b) is

$$R(A,b) := \sup\{r \geq 0 \,|\, K_A(-r), J_A(-r), K_b(-r), J_b(-r) \geq 0\}$$

(the supremum of the empty set is taken $-\infty$).

Remark 2. It can be proven (see [11]) that in case $R(A,b) > 0$ the scheme functions are absolutely monotonic, consequently non-negative on $[-R(A,b), 0]$. Hence, by $J_b(0) = b^T$ and $J_A(0) = A$, $A, b \geq 0$ is necessary for $R(A,b) > 0$. This condition is violated by many methods that are used in practice (see e.g. [3]). Moreover, $R(A,b) = +\infty$ only for some first order methods, for example for the implicit Euler method.

5 Discrete positive invariance of RK methods

Now we are in the position to formulate the main goal of the paper precisely as follows:

Let Φ, $\mathcal{C}$, RK(A,b) be given as above and suppose that $\mathcal{C}$ is positively invariant w.r.t. Φ.
Construct $H = H(f;\mathcal{C};A,b)$ directly such that $\mathcal{C}$ is discrete positive invariant under RK(A,b) with step size constant H.

Remark 3. This problem is considered in the literature by several authors for the classical problem of non-negativity conservation, i.e. discrete positive invariance of the positive orthant $[0,\infty)^N$, see e.g. [2], [5], [10], [6], [9] and the references therein and for general cones in [7] and [8]. One of the conclusions of these papers (see [2], [9]) is that such an $H > 0$ does *not* exist for all Φ, $\mathcal{C}$ and RK(A,b), see [9] for counterexamples.

Theorem 1. *Let f, Φ as above, $\mathcal{C}$ a convex, closed polyhedral set with non-empty interior and $\varepsilon > 0$ such that (8) holds. Further let RK(A,b) be given with $R(A,b) > 0$.*

Then $\mathcal{C}$ is discrete positive invariant under RK(A,b) with step size constant H where

$$H = \min\{\varepsilon R(A,b),\ H_{\text{def}}\}. \tag{12}$$

Proof. Let $u_0 \in \mathcal{C}$, $h \in [0, H_{\mathrm{def}}]$. We consider the first step $(n = 0)$ as

$$u_1 = u_0 + h(b^T \otimes I)F(y) \tag{13}$$
$$y = (e \otimes I)u_0 + h(A \otimes I)F(y) \tag{14}$$

with

$$y := \begin{pmatrix} y_1 \\ \vdots \\ y_s \end{pmatrix},\ F(y) := \begin{pmatrix} f(y_1) \\ \vdots \\ f(y_s) \end{pmatrix}.$$

Note that here $\otimes$ denotes the Kronecker product of matrices, i.e. if $Q \in \mathbb{R}^{q_1 \times q_2}$ and $R \in \mathbb{R}^{r_1 \times r_2}$ are arbitrary matrices then $Q \otimes R$ is the real $q_1 r_1$-by-$q_2 r_2$ matrix with block structure $Q \otimes R = \begin{pmatrix} Q_{1,1}R & \ldots & Q_{1,q_2}R \\ \vdots & & \vdots \\ Q_{q_1,1}R & \ldots & Q_{q_1,q_2}R \end{pmatrix}$.

One of the main idea of the proof is to write f in a quasy-linear form: $f(v) = -\frac{1}{\varepsilon}v + \frac{1}{\varepsilon}p(v)$ with $p(v) := v + \varepsilon f(v)$ and $p(v) \in \mathcal{C}$ whenever $v \in \mathcal{C}$. Then, making use of the notation $P(y) := (p(y_1), ..., p(y_s))^T$, we have $F(y) = -\frac{1}{\varepsilon}y + \frac{1}{\varepsilon}P(y)$ and $P(y) \in \mathcal{C}^s$ whenever $y \in \mathcal{C}^s$. Hence (14) reads

$$y = (e \otimes I)u_0 + (A \otimes I)(-\frac{h}{\varepsilon}y + \frac{h}{\varepsilon}P(y))$$

and, solving it formally to y we obtain by the term $\mathbb{I} := I \otimes I$ with identity matrices of order $s \times s$ and $N \times N$, respectively

$$y = \underbrace{\left(\mathbb{I} + \frac{h}{\varepsilon}(A \otimes I)\right)^{-1}(e \otimes I)}_{\left((I + \frac{h}{\varepsilon}A)^{-1}e\right) \otimes I} u_0 + \frac{h}{\varepsilon} \underbrace{(A \otimes I)\left(\mathbb{I} + \frac{h}{\varepsilon}(A \otimes I)\right)^{-1}(e \otimes I)}_{\left(A(I + \frac{h}{\varepsilon}A)^{-1}\right) \otimes I} P(y).$$

Substituting this formula into (13) and writing out the components of y and using the notation of the scheme functions introduced in the previous section we arrive at

$$y_i = K_A(-\frac{h}{\varepsilon})_i + u_0 + \sum_j \frac{h}{\varepsilon} J_A(-\frac{h}{\varepsilon})_{ij} p(y_j) \qquad i = 1, ..., s \tag{15}$$

$$u_1 = K_b(-\frac{h}{\varepsilon})_i u_0 + \sum_j \frac{h}{\varepsilon} J_b(-\frac{h}{\varepsilon})_{ij} p(y_j). \tag{16}$$

Moreover,

$$K_A(-\frac{h}{\varepsilon})_i + \sum_j \frac{h}{\varepsilon} J_A(-\frac{h}{\varepsilon})_{ij} = 1$$

and

$$K_b(-\frac{h}{\varepsilon}) + \sum_j \frac{h}{\varepsilon} J_b(-\frac{h}{\varepsilon})_j = 1.$$

Thus, in (15) and (16) y_i and u_1 are written as convex combinations of u_0 and $p(y_j)$ which are elements of $\mathcal{C}$ whenever $\frac{h}{\varepsilon} \leq R(A,b)$ and $y_j \in \mathcal{C}$, respectively. Finally, application of some continuation techniques (e.g. almost word-by-word application that of Theorem 1 in [6]) results in the statement of the theorem. □

Remark 4. From the constructions in [9] it can be shown that H in (12) is the largest possible constant for positive invariance of $\mathcal{C} = [0,\infty)^3$ for some functions f whenever RK(A,b) is irreducible and non-confluent.

Corollary 1. *Applying Theorem 1 we arrived at a quite simple formula, essentially $h \leq \varepsilon R(A,b)$ for step sizes that guarantee discrete positive invariance of $\mathcal{C}$ under RK(A,b). This gives us a practically useful means to determine* a priori *the step sizes to be chosen for RK methods to possess the invariance property: ε can be determined, for example, by applying Lemma 3 and $R(A,b)$ by its definition.*

Example 1. Let us consider the heat conduction problem and its mathematical modeling as is given in Section 1 above with equidistant space grid. The resulting ODE reads $U' = LU$ with $L = \text{tridiag}(a_k, b_k, a_{k+1})$ where $a_k = \sigma(x_k - \Delta x/2)/(\Delta x)^2$, $b_k = -(a_k + a_{k+1})$, $\Delta x = \omega/(N+1)$, $x_k = k\Delta x$ (for all k). Consider $\mathcal{C} = [0,\eta]^N$ with some $\eta > 0$ constant. Then, using the notations of the paper, $\mathcal{C}$ is the intersection of $K = 2N$ half-spaces with $\varphi_i = e_i$, $\varphi_{N+i} = -e_i$ and $p_i = 0$, $p_{N+i} = -\eta$ ($i = 1, ..., N$ and e_i is the ith unit vector of $\mathbb{R}^N$). Simple calculation shows that condition (8) holds with $\ell = \min_k(\sigma(x_k - \Delta x/2) + \sigma(x_k + \Delta x/2))/(\Delta x)^2$. Further, for the explicit Euler, the implicit trapesoid rule (Crank-Nicolson method) and the implicit Euler method the absolute monotonicity radius equals 1, 2, ∞, respectively. Hence these methods preserve the positive invariance of $\mathcal{C}$ with step sizes $h \leq H$ where H equals $1/\ell$, $2/\ell$ and ∞, respectively. Observe that this condition is very severe for the first two methods because $\varepsilon = 1/\ell$ is proportional to $(\Delta x)^2$ hence it is in practical situations (when N is large) very small. In contrast to these methods, the implicit Euler method lets $\mathcal{C}$ invariant for all positive step sizes.

References

[1] Blanchini, F., Set invariance in control. *Automatica* **35**, 11 (1999) 1747–1767.

[2] Bolley, C. , Crouzeix, M., Conservation de la positivité lors de la discrétisation des problèmes d'évolution paraboliques, *R.A.I.R.O. Analyse numérique* **12** 3 (1978) 237–245.

[3] BUTCHER, J. C., *Numerical methods for ordinary differential equations.* Chichester: Wiley (2003).

[4] DEIMLING, K., *Ordinary differential equations in Banach spaces.* Lecture Notes in Mathematics. 596. Berlin-Heidelberg-New York: Springer-Verlag. VI (1977)

[5] GERISCH, A., WEINER, R., The positivity of low order explicit Runge-Kutta schemes applied in splitting methods. *Comp. and Math. with Appl.* **45** (2003), 53–67

[6] HORVÁTH, Z., Positivity of Runge-Kutta and diagonally split Runge-Kutta methods. *Applied Numerical Mathematics*, **28**, 2–4 (1998) 309–326.

[7] HORVÁTH, Z., On the positivity of matrix-vector products. *Linear Algebra and its Applications.* **393C** (2004) 253–258 .

[8] HORVÁTH, Z., Positively invariant cones of dynamical systems under Runge-Kutta and Rosenbrock-type discretization. *PAMM* (*Proceedings in Applied Mathematics and Mechanics*) **4**, 1, (2004) 688*-689.

[9] HORVÁTH, Z., On the positivity step size threshold of Runge-Kutta methods. *Applied Numerical Mathematics*, **53** (2005) 341–356.

[10] HUNDSDORFER, W., VERWER, J., *Numerical solution of time-dependent advection-diffusion-reaction equations.* Springer Series in Computational Mathematics. 33. Berlin: Springer, 2003

[11] KRAAIJEVANGER, J. F. B. M., Contractivity of Runge-Kutta methods, *BIT* **31** (1991) 482–528.

[12] MARTIN, R. H. JR., *Nonlinear operators and differential equations in Banach spaces.* Pure and Applied Mathematics. New York etc.: John Wiley&Sons, a Wiley-Interscience Publication (1976).

Investigations of Gyula Farkas in the Fields of Electrodynamics and Relativity Theory

Zoltán GÁBOS* and László NAGY**

Faculty of Physics, Babeş-Bolyai University
str. Kogălniceanu nr. 1
RO 400084 Cluj, Romania

Abstract. Gyula Farkas was an outstanding representative of the phenomenological physics. He has followed continuously the evolution of the theory of electromagnetic phenomena. He has obtained notable results related to the electromagnetic interaction of current elements, the hidden relationships between the equations of Maxwell's theory, and the consequences of the Lorentz's electron theory for the theory of the continuum.

He was between the few physicists, who have realized the importance of the theory of relativity in the time of its founding. He was well informed both about the results of H.A. Lorentz and A. Einstein. We emphasize two of his results: he has obtained formulae, equivalent with the Lorentz transformations, in an original way; in the case of the electromagnetic field in the vacuum, he has given the transformation formulae, different from the usual ones. It follows from these transformation formulae, that there exists a reference system K_F moving with velocity $\mathbf{n}$ relative to K, for which the observers from K and K' give the same time and the same $\mathbf{E}$ and $\mathbf{B}$ field vectors.

AMS 2000 subject classifications: 01A60, 83A05
Key words and phrases: Gyula Farkas, electrodynamics, relativity theory

The university education in Cluj began with the foundation of the Jesuit college in 1581 by Stephanus Báthory, the prince of Transylvania and king of Poland. However, university education was not continuous in the following period.

The foundation of the modern university in Cluj was done in 1872, which in a few years has reached European level. Many outstanding Hungarian scientists have been appointed as professors at the university, one of them being Gyula Farkas. He came to the university in 1887, and served it for 28 years. He was equally interested in mathematics and physics.

Before coming to Cluj, he has studied mainly mathematics. He has achieved recognized results in algebra, function theory and geometry. In 1881 he has obtained his doctoral degree in mathematics as principal topic and in natural sciences and astronomy as secondary topics.

After Gyula Farkas was appointed professor of theoretical physics at the university of Cluj, he proved to be equally interested in physics, too. Between 1887 and 1892 he became familiar with the last achievements in physics. Beginning

* email: zgabos@phys.ubbcluj.ro
** email: lnagy@phys.ubbcluj.ro

with 1893 he started his own, original research. After finding the research field suitable for his abilities, he has enriched the theoretical physics with many new results. In the followings we will show his results in the fields of electrodynamics and relativity theory.

1 The first investigations

In the first years of his activity at the University of Cluj was decisive for him the collaboration with the famous chemistry professor of the university, Rudolf Fabinyi. He has approached theoretical physics through chemical physics. Farkas was strongly influenced by the book of W. Ostwald, "Lehrbuch der allgemeinen Chemie". His investigations concerning the theory of the galvanic elements brought his attention to the electric current and to the transformation of energy [1,2,3]. His ideas about the concept of energy reappeared also later, emphasized also by the choice of the title of his electrodynamics lecture notes appeared in 1908 and 1913, "The propagation of the energy".

Farkas has published his results concerning the electric currents in his lecture notes appeared in 1890. He has closed his investigations in this direction with an article published in 1893 [4].

Ampère gave a law for the interaction between the elementary currents $Id\mathbf{s}$ and $I'd\mathbf{s}'$, where I and I' stand for the electric current intensities. Gyula Farkas has studied the interaction between the electric current in a closed circuit and the elementary current $I'd\mathbf{s}'$. The force is given by the vectorial sum of the elementary interactions. It is possible, that different elementary laws lead to the same global result. Ampère has formulated the interaction law in two equivalent forms from the point of view of the global effect. In his paper, Gyula Farkas has given a method for obtaining all the equivalent elementary interaction laws.

The investigations related to electricity and magnetism made possible to Gyula Farkas to find the way to the electrodynamics of Maxwell.

2 The investigations of Gyula Farkas in electrodynamics

A turning point in the scientific activity of Gyula Farkas was brought by an honorable charge. He has represented the University of Cluj and the Mathematical and Physical Society at a celebration in December 1892, at the University of Padova, on the occasion of 300 years from Galileo's inauguration as a chair. As a part of the celebration he has received the Doctor Honoris Causa degree of the University of Padova. Preparing to these celebration he got notice of Galileo's results on forced constrained motions. This problem became one of his main interests in all of his activity. But it was also decisive that in Padova he has met personally the famous professor from Göttingen, W. Voigt, who was one of the most outstanding representative of the phenomenological school. Gyula Farkas became one of the most consequent representative of this trend. In the second part of the 19th century one of the most important result of the phenomenological physics was the electrodynamics of Maxwell. This new theory came rapidly into the attention of Farkas.

Probably in the study of Maxwell's theory Farkas was based on the two volumes monograph by W. Voigt, "Kompendium der theoretischen Physik" published in 1895 and 1896, and on the monograph by C. Cristiansen, "Elemente der theoretischen Physik" published in 1894. He has written reviews on both of these books.

He became very familiar with this new theory also because of his special abilities and character.

Farkas regarded mathematics as one of the most excellent tool of physics. As Galilei, he also declared "Nature speaks to us in the language of mathematics". He knew very well this language, and if necessary, he was able also to enrich it. He was one of the most outstanding Hungarian experts in vector analysis. His only printed book, "Vector and simple inequalities theory", published in 1900, and one of his papers [5] represent this field. Therefore he had no mathematical problems at all regarding Maxwell's theory, and this fact helped him to discover the great importance of this new theory.

The new theories were always a challenge for him. He always was up to date with the new publications, he received very quickly the new theories and built them into the curriculum for students. He considered this the basic requirement for a conscientious career. He always emphasized, that even the teaching of classical knowledge should be done only being up to date with the newest achievements in science.

He has taught the electromagnetic theory on the level of his time, being an example to other universities, which were more conservative. He accentuated that the electromagnetic interaction has not an action at a distance character. He has said, that "If somewhere electricity, magnetism or electric current occurs, its influence, according to the theory, does not appear instantly at distance, but it propagates with a finite velocity in all directions".

In his time most of the physicists accepted that the medium for the electromagnetic influences is a special matter called "aether", and Farkas had adopted this idea. At September 22nd, 1907, in his speech on the occasion of his appointment as the rector of the University, he has said: "... the electric and magnetic influences can be attributed to the tensions and to the transversal forces appearing in the bodies and in the aether, the weightless matter filling the space between bodies, and with the propagation of these tensions from particle to particle with a finite velocity, also the electric and magnetic influences will propagate with this finite velocity in the space." And also later, in his lecture notes published in 1914 entitled "Analytical mechanics", he has written about the aether in rest, which fills the Universe till the infinity.

Adopting the existence of the aether withheld him to accept the electromagnetic field. He regarded only the electric charge as physical reality, and not the electromagnetic field. As a consequence, for him the basic equation of electromagnetism was the continuity equation expressing charge conservation. It is well known, that from the Maxwell equations

$$\text{rot } \mathbf{H} = -\frac{\partial \mathbf{D}}{\partial t} + \mathbf{j_e}, \qquad \text{div } \mathbf{D} = \rho_e \tag{1}$$

one can derive the continuity equation

$$\frac{\partial \rho_e}{\partial t} + \text{div } \mathbf{j_e} = 0. \tag{2}$$

Gyula Farkas has regarded the equations (1) as the solutions expressed with vector parameters of equation (2).

Farkas also found very useful the notions of magnetic charge density and magnetic current density. Starting from the continuity equation

$$\frac{\partial \rho_m}{\partial t} + \text{div } \mathbf{j_m} = 0 \tag{3}$$

he has derived

$$\text{rot } \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} - \mathbf{j_m}, \qquad \text{div } \mathbf{B} = \rho_m. \tag{4}$$

After this, he had eliminated the quantities ρ_m and $\mathbf{j_m}$, obtaining the well known Maxwell equations. It is interesting, that nowadays the physicists looking for the magnetic monopoles are using the extended Maxwell equations similar to (4).

3 The investigations of Gyula Farkas in relativity theory

Gyula Farkas has followed in detail the evolution of electrodynamics and of the electron theory of H. A. Lorentz, consequently he has recognized the importance of the special relativity theory in the first years after its publication.

The works and the ideas of Lorentz helped him to approach to the relativity theory. But the same ideas of Lorentz blocked him to reach the ideas of Einstein.

In his two papers published in Physikalische Zeitschrift in 1906 and 1907 [7,8] he used the aether-hypothesis and the framework set up by Lorentz. Later he has adopted an intermediary position, and in his papers written in the concept of Lorentz has recognized the ideas of Einstein as a possible alternative. In his paper published in 1915 he has written "The system of Einstein is self consistent, and till now there are no experimental observations, with which the 'relativity theory' would not be compatible".

We can be certain, that he knew the articles of Einstein. This is attested by the notes of Farkas, which we have found on the side of Einstein's article published in the 4th volume of Jahrbuch der Electronik und Radioactivität.

Gyula Farkas assumed a pioneer role, when since 1910 he has taken into account the requirements of the relativity theory in the mechanics of elastic bodies and ideal fluids, and in the electrodynamics of continuous media. The relativistic generalization of virtual change made possible to him to initiate the relativistic treatment of constrained motions [9]. In his lecture notes published in 1908 entitled "The propagation of the energy" he has included also the relativity theory, in the time when most of the physicists did not understand yet the new theory. In 1922 he has attempted to explain gravitation within the special relativity theory [11].

In the next part we present one of his results. Farkas derived some of the transformation formulae of the relativity theory in an unusual way.

In establishing the Lorentz-Herglotz transformation formulae, expressing the relationship between space and time, he had three basic assumptions:

1. The transformation is the mapping of spacetime systems.
2. There is a velocity, which has the same magnitude in both spacetime systems, regardless of position, time or direction.
3. The transformation is finite.

In case of the second assumption he has thought to the velocity of light in vacuum, in accordance with Einstein. In one of his papers he mentions, that the assumptions No 1 and 2 were also formulated by C. Munari in the 23rd volume of Rendiconti del Reale Academia dei Lincei [10].

Based on the three basic assumptions above, Gyula Farkas has obtained the following transformation formulae

$$\mathbf{x}' + \mathbf{n}t' = \mathbf{x} - \mathbf{n}t \tag{5}$$

$$t' + \frac{1}{c^2}(\mathbf{n}\cdot\mathbf{x}') = t - \frac{1}{c^2}(\mathbf{n}\cdot\mathbf{x}), \tag{6}$$

where

$$\mathbf{n} = \frac{\gamma}{1+\gamma}\mathbf{v}; \qquad \gamma = \left(1 - \frac{v^2}{c^2}\right)^{-\frac{1}{2}}. \tag{7}$$

In the case of the electromagnetic field in the vacuum, he has given the transformation formulae, different from the usual ones

$$\mathbf{B}' + \frac{1}{c^2}\mathbf{n}\times\mathbf{E}' = \mathbf{B} - \frac{1}{c^2}\mathbf{n}\times\mathbf{E} \tag{8}$$

$$\mathbf{E}' - \mathbf{n}\times\mathbf{B}' = \mathbf{E} + \mathbf{n}\times\mathbf{B}. \tag{9}$$

$>$From the above one may obtain easily the usual transformation formulae. Let us make the scalar product of both sides of Equation 5 by **n**, obtaining

$$(\mathbf{n}\cdot\mathbf{x}') + t'\mathbf{n}^2 = (\mathbf{n}\cdot\mathbf{x}) - t\mathbf{n}^2.$$

Using this relationship and (6), we get

$$t' = \gamma\left[t - \frac{1}{c^2}(\mathbf{v}\cdot\mathbf{x})\right]. \tag{10}$$

Based on (5) and (10) one obtains

$$\mathbf{x}' = \mathbf{x} + (\gamma - 1)(\mathbf{e}\cdot\mathbf{x})\mathbf{e} - \gamma v t \mathbf{e}, \tag{11}$$

where

$$\mathbf{e} = \frac{\mathbf{v}}{v}. \tag{12}$$

Starting from (8) and (9) we can get also the usual formulae. Let us multiply vectorially (8) from left with **e**, taking into account (9) and the relationship obtained from it

$$(\mathbf{e}\cdot\mathbf{E}') = (\mathbf{e}\cdot\mathbf{E}). \tag{13}$$

By this procedure we arrive to the well known transformation formula

$$\mathbf{E}' = \gamma\mathbf{E} + (1-\gamma)(\mathbf{e}\cdot\mathbf{E})\mathbf{e} + \gamma v(\mathbf{e}\times\mathbf{B}). \tag{14}$$

If (9) is multiplied from left by **e**, and we take into account (8) and the relationship obtained from it

$$(\mathbf{e}\cdot\mathbf{B}') = (\mathbf{e}\cdot\mathbf{B}), \tag{15}$$

we get the usual transformation formula for the magnetic field

$$\mathbf{B}' = \gamma\mathbf{B} + (1-\gamma)(\mathbf{e}\cdot\mathbf{B})\mathbf{e} - \gamma\frac{v}{c^2}(\mathbf{e}\times\mathbf{E}) \tag{16}$$

For the reversed formulae we can write

$$\mathbf{E} = \gamma\mathbf{E}' + (1-\gamma)(\mathbf{e}\cdot\mathbf{E}')\mathbf{e} - \gamma v(\mathbf{e}\times\mathbf{B}') \tag{17}$$

$$\mathbf{B} = \gamma\mathbf{B}' + (1-\gamma)(\mathbf{e}\cdot\mathbf{B}')\mathbf{e} + \gamma\frac{v}{c^2}(\mathbf{e}\times\mathbf{E}'). \tag{18}$$

4 The deduction of the Lorentz-Farkas transformation formulae

Because the lecture notes of Farkas, which contains the detailed proof of is formulae, to our knowledge, are lost, we give our own deduction.

We assume, that the perpendicular component of the position vector **x** relative to the velocity vector **v**, does not change. Then

$$(\mathbf{v}\times\mathbf{x}') = (\mathbf{v}\times\mathbf{x}), \tag{19}$$

so we can write

$$v^2\mathbf{x}'^2 - (\mathbf{v}\cdot\mathbf{x}')^2 = v^2\mathbf{x}^2 - (\mathbf{v}\cdot\mathbf{x})^2. \tag{20}$$

We are looking for a linear transformation with the aid of the scalar quantities t, t', $(\mathbf{v}\cdot\mathbf{x})$, $(\mathbf{v}\cdot\mathbf{x}')$, and the vector quantities $\mathbf{x}$, $\mathbf{x}'$ and $\mathbf{v}$

$$a_1\mathbf{x}' + a_2t'\mathbf{v} = a_3\mathbf{x} + a_4t\mathbf{v} \tag{21}$$

$$b_1t' + b_2(\mathbf{v}\cdot\mathbf{x}') = b_3t + b_4(\mathbf{v}\cdot\mathbf{x}). \tag{22}$$

If the coefficients are functions of v^2, for the reversed transformation we can write

$$a_1\mathbf{x} - a_2t\mathbf{v} = a_3\mathbf{x}' - a_4t'\mathbf{v} \tag{23}$$

$$b_1t - b_2(\mathbf{v}\cdot\mathbf{x}) = b_3t' - b_4(\mathbf{v}\cdot\mathbf{x}'). \tag{24}$$

Comparing the direct and the reversed relationships, one obtains the equalities

$$a_3 = a_1, \quad a_4 = -a_2, \quad b_3 = b_1, \quad b_4 = -b_2. \tag{25}$$

Using these, and dividing (21) and (22) by a_1 and b_1, respectively, we obtain

$$\mathbf{x}' + At'\mathbf{v} = \mathbf{x} - At\mathbf{v} \tag{26}$$

$$t' + B(\mathbf{v}\cdot\mathbf{x}') = t - B(\mathbf{v}\cdot\mathbf{x}). \tag{27}$$

Using the second basic assumption we obtain

$$\mathbf{x}'^2 - c^2t'^2 = 0; \quad \mathbf{x}^2 - c^2t^2 = 0. \tag{28}$$

Using (26), (27) and (28) we can prove that $B = A/c^2$. As a consequence, (27) may be replaced by

$$t' + \frac{A}{c^2}(\mathbf{v}\cdot\mathbf{x}') = t - \frac{A}{c^2}(\mathbf{v}\cdot\mathbf{x}). \tag{29}$$

For the origin of system K' one can write

$$\mathbf{x}' = 0 \qquad \mathbf{x} = \mathbf{v}t.$$

Using (26) and (29) we obtain for A the equation

$$\frac{v^2}{c^2}A^2 - 2A + 1 = 0,$$

which has the solutions

$$A_1 = \frac{\gamma}{\gamma+1}; \qquad A_2 = \frac{\gamma}{\gamma-1}.$$

Because A_2 is not finite for $v = 0$, according to the third basic assumption only the solution A_1 is acceptable. With this value from (26) and (29) we obtain the transformation formulae (5) and (6).

In the followings we make two observations concerning the derived transformation formulae, which we may call the Lorentz-Farkas formulae.

Let us consider a system K_F which moves relative to K with velocity

$$\mathbf{v_F} = \mathbf{n} = \frac{\gamma}{\gamma+1}\mathbf{v}, \tag{30}$$

which has the direction of $\mathbf{v}$. For an observer from K', the system K_F moves with

$$v'_F = \frac{-v + v_F}{1 - \frac{vv_F}{c^2}} = -v_F, \tag{31}$$

along the same direction given by $\mathbf{v}$.

For the origin O_F of the system K_F, the observers of the systems K and K' may write

$$\mathbf{x_F} = \mathbf{v_F}t; \qquad \mathbf{x'_F} = -\mathbf{v_F}t', \tag{32}$$

respectively. Taking into account (32) and (6), for the origin O_F one obtains

$$t'_F = t_F. \tag{33}$$

This special reference system K_F, which has a special symmetry relative to the systems K and K' may be called the Farkas reference system.

Our second observation concerns the transformation formula (6). Let us consider the components of the position and velocity four-vectors, respectively

$$\{x_1, x_2, x_3, ict\}; \quad \{\gamma v_1, \gamma v_2, \gamma v_3, ic\gamma\}.$$

Using these, we may construct the invariant scalar

$$-c^2\left[t - \frac{1}{c^2}(\mathbf{x}\cdot\mathbf{x})\right]. \tag{34}$$

Including in this expression the data $\mathbf{x}, t, \mathbf{v} = \mathbf{v_F}$ for system K and $\mathbf{x}', t', \mathbf{v} = -\mathbf{v_F}$ for system K', we obtain easily the transformation formula (6).

5 Transformation formulae for the electromagnetic field

In this section we will derive the transformation formulae (8) and (9) given by Gyula Farkas.

The components of the electromagnetic field tensor in vacuum may be written as

$$\begin{gathered} F_{12} = -F_{21} = B_3;\; F_{23} = -F_{32} = B_1;\; F_{31} = -F_{13} = B_2 \\ F_{41} = -F_{14} = \frac{i}{c}E_1; F_{42} = -F_{24} = \tfrac{i}{c}E_2; F_{43} = -F_{34} = \frac{i}{c}E_3. \end{gathered} \tag{35}$$

Using these quantities and the velocity four-vector components denoted by u_ν, we obtain

$$-\frac{i}{2c}\epsilon_{j\nu\rho\sigma}u_\nu F_{\rho\sigma} = \gamma\left[B_j - \frac{1}{c^2}(\mathbf{v}\times\mathbf{E})_j\right] \tag{36}$$

$$F_{j\nu}u_\nu = \gamma[E_j + (\mathbf{v}\times\mathbf{B})_j]. \tag{37}$$

Using the same procedure as in the previous section, one may obtain from equations (36) and (37) the transformation formulae (8) and (9), respectively.

Based on (14) and (16), for the system K_F diverging from K with velocity $\mathbf{v_F}$, we obtain

$$\mathbf{E_F} = \gamma_F\mathbf{E} + (1-\gamma_F)(\mathbf{e}\cdot\mathbf{E})\mathbf{e} + \gamma_F v_F(\mathbf{e}\times\mathbf{B}) \tag{38}$$

$$\mathbf{B_F} = \gamma_F\mathbf{B} + (1-\gamma_F)(\mathbf{e}\cdot\mathbf{B})\mathbf{e} - \gamma_F\frac{v_F}{c^2}(\mathbf{e}\times\mathbf{E}), \tag{39}$$

where

$$\gamma_F = \left(1 - \frac{v_F^2}{c^2}\right)^{-\frac{1}{2}} = \sqrt{\frac{1+\gamma}{2}} \tag{40}$$

The reversed transformation, from K_F to K', based on (17) and (18), is obtained

$$\mathbf{E'_F} = \gamma_F\mathbf{E'} + (1-\gamma_F)(\mathbf{e}\cdot\mathbf{E'})\mathbf{e} - \gamma_F v_F(\mathbf{e}\times\mathbf{B'}) \tag{41}$$

$$\mathbf{B'_F} = \gamma_F\mathbf{B'} + (1-\gamma_F)(\mathbf{e}\cdot\mathbf{B'})\mathbf{e} + \gamma_F\frac{v_F}{c^2}(\mathbf{e}\times\mathbf{E'}). \tag{42}$$

Using (41) and (14), then (42) and (16), we obtain the following relations, which expresses the symmetry of systems K and K' relative to K_F

$$\mathbf{E'_F} = \mathbf{E_F}; \quad \mathbf{B'_F} = \mathbf{B_F}. \tag{43}$$

Here again, we observe the special symmetry features of the Farkas reference system.

6 Conclusions

Gyula Farkas was one of the outstanding scientists of the Hungarian university in Cluj, founded in 1872. In this paper we presented some results of Farkas in electrodynamics and relativity theory. He has written his relativistic transformation formulae using a reference system (which we may call Farkas-type), symmetric relative to the two reference systems involved in the transformation.

References

[1] Farkas, Gy. and Fabinyi, R., Constant electric current by the oxidation of carbon, *Vegytani Lapok* **5** (1887) 97–112.

[2] Farkas, Gy., On the connections of chemical and electric energy, *Erdélyi Múzeumegylet Értesítője* **13** (1888) 33–38.

[3] Farkas, Gy., Remarks on the theory of steady and constant electric current, *Erdélyi Múzeumegylet Értesítője* **17** (1892) 165–240.

[4] Farkas, Gy., The determination of the equivalents of Ampère's elementary laws, *Értekezések a mathematikai tudomńyok köréből* **15** (1893) 1–50.

[5] Farkas, Gy., Complements to vector-theory and the theory of elecromagnetism, *Matematikai és természettudományi értesítő* **16** (1898) 321–360.

[6] Farkas, Gy., General mechanical principles of the ether, *Matematikai és természettudományi értesítő* **19** (1901) 99–102.

[7] Farkas, Gy., Über den Einfluss der Erdbewegung auf electromagnetische Erscheinungen, *Physicalische Zeitschrift* **7** (1906) 654–657.

[8] Farkas, Gy., Über das Postulat der Relativität, *Physicalische Zeitschrift* **8** (1907) 169–171.

[9] Farkas, Gy., Foundation of the continuity theory of electricity and magnetism, *Matematikai és természettudományi értesítő* **28** (1910) 1–25, **29** (1911) 777–809.

[10] Farkas, Gy., Comments on negative experiment of Michelson, *Matematikai és természettudományi értesítő* **33** (1915) 355–361.

[11] Farkas, Gy., The Einstein's gravitation from the old theory, *Matematikai és természettudományi értesítő* **39** (1922).

Proof of the existence of the integrating factor based on the Farkas Lemma

Katalin MARTINÁS*, Fabian BODOKY**, and Ildikó BRODSZKY***

[1] Department of Atomic Physics, Loránd Eötvös University
Pázmány Péter sétány 1, H-1117 Budapest, Hungary
[2] Institut für Physik, Universität Basel
Klingelbergstrasse 82, CH-4056 Basel, Switzerland
[3] AUTOKUT RT, Csóka utca 7, H-1115 Budapest, Hungary

Abstract. Gyula Farkas, professor for theoretical physics in Kolozsvár at the end of the nineteenth century, developed a mathematically rigorous formalization of thermodynamics. Starting from the First and Second Law, as well as from the Farkas Lemma of thermodynamics, he proved the existence of entropy and of a universal temperature scale. His original paper is very terse, this is the first elaborate proof.

AMS 2000 subject classifications: 80A06, 01A06
Key words and phrases: thermodynamics, entropy, Gyula Farkas

1 Introduction

Gyula Farkas (1874 - 1930), sometimes called Julius Farkas in German publications, was a professor for theoretical physics at Kolozsvár when he published a paper with a mathematically exact introduction of absolute entropy and temperature in 1886, founded on a lemma and a theorem:

> *Farkas Lemma of Thermodynamics*: In reversible processes, no body (or system of bodies) can go adiabatically into a state to which it can go by pure heat exchange, i.e. by changing only the temperature by supplying or abstracting heat.

As a consequence, Farkas proved the

> *Farkas Theorem of Thermodynamics*: In reversible processes, heat elements absorbed by a body (or a system of bodies) always have integrating divisors, and one of them is an identical function of the empirical temperature $\mathrm{d}S = \mathrm{d}'Q/T$, i.e. there is an absolute entropy and an absolute temperature scale (up to a constant factor).

* email: martinas@elte.hu
** email: f.bodoky@tnw.tudelft.nl
*** email: brcica@netquick.hu

Farkas did his work not only earlier than Carathéodory, but he proved the *global* existence of the absolute temperature as integrating factor (using the Lebesgue lemma), where Carathéodory did only a *local* proof. Farkas' proof remained unnoticed, perhaps for its extraordinary terseness; so the aim of this paper is to present the results of Farkas with a comprehensive and elaborate proof.
We title the theorem the *Farkas theorem of thermodynamics*, since there is another lemma by Farkas (which is nowadays known as the Farkas lemma); it is about linear inequalities and mainly used for linear programming [1, 2].

2 Historical Remarks

Rudolf Clausius was the first to formulate the integral form of the Second Law[3], but he based his reasoning on physical considerations (i.e. the Carnot process), which could not be connected convincingly with mathematics. His main merit was to give the law a clear mathematical form, namely that absolute temperature is an integrating factor of heat

$$\mathrm{d}S = \mathrm{d}'Q/T \tag{1}$$

This inspired Gustav Zeuner to look for an integrating factor to identify it with the absolute temperature[4]. He has shown that the existence of entropy is a consequence of the First Law. He had two main aims:

1. To deduce both the First and the Second Law from the principle of the equivalence of heat and work without using any other principle.
2. Not to use the absolute temperature as a primary concept at the beginning, rather to define it as a result.

But Zeuner - as well as Clausius, who objected strongly against Zeuner's results - did not investigate whether a solution exists: In fact, the existence is trivial for two variables, but not for more. Woldemar Voigt, professor at Göttingen, examined the expression of elementary heat in the case of n variables[5]. He realized that the existence of an integrating factor for heat is mathematically equivalent with the existence of $n-1$ dimensional adiabatic surfaces (which are the geometrical places of all those states that are adiabatically accessible from a given state). Gyula Farkas noticed that Voigt only *assumed* the existence of these surfaces, and supplied mathematical reasoning for the existence of these surfaces, which Voigt referred to in the second edition of his book.
The other person connected to a mathematical deduction of thermodynamics is Constantin Carathéodory[6]. It is interesting to remark that Caratheéodory was a student of Max Born, who was a student of Voigt; so we might speculate whether Carathéodory knew about Farkas' article. Carathéodory's postulate system is one of the most elegant constructions of thermodynamics. It shows that the so called *adiabatic inaccessibility postulate* is sufficient to ensure the existence of an integrating factor for the heat element of the First

Law independently of the number of variables, i.e. $\mathrm{d}'Q$ can always be written $\mathrm{d}'Q = t\,\mathrm{d}s$, where t and s are some (yet unspecified) functions. Only postulating the existence of thermal equilibrium and introducing the concept of empirical temperature, it can be shown that one of the *possible* functions depends only on temperature. The deficiency of Carathéodory's approach, however, is that it merely proofs the *local* existence of entropy and absolute temperature[7].
Carathéodory's postulate system is a deductive system, i.e. a general theorem is postulated and its consequences tested directly - the theorem being the adiabatic inaccessibility principle for processes (meaning that in whatever small neighborhood of a state there are states which are inaccessible by adiabatic functions).
One of the most fascinating aspects of Carathéodory's construction is that he always makes a very clear distinction between mathematical principles and pieces of physical experience. But the main difference to Farkas' approach is that Carathéodory built a mathematical formalism and looked for a physical process to apply it to, whereas Farkas started with the First and Second Law and deduced their mathematical consequences. For a more comprehensive and historical overview see Refs. [8] and [9].

3 Mathematical Tools

The main mathematical knowledge needed is about Pfaffian expressions. A Pfaffian form is a special case of a k-form, namely a 1-form. A differential equation is called a Pfaffian expression, if it has the following form:

$$\begin{aligned}\mathrm{d}'w = {} & A_1\,(x_1, x_2, \ldots, x_n)\;\mathrm{d}x_1 \;+\; A_2\,(x_1, x_2, \ldots, x_n)\;\mathrm{d}x_2 \\ & + \;\ldots\; + \;A_n\,(x_1, x_2, \ldots, x_n)\;\mathrm{d}x_n. \end{aligned} \tag{2}$$

If $\mathrm{d}w$ is a so-called total (or exact) differential, the integral of $\mathrm{d}w = 0$ is independent of the path and the Pfaffian equation has a constant solution

$$\eta\,(x_1, \ldots, x_n) = c. \tag{3}$$

If $\mathrm{d}'w$ is not a total differential (the prime indicating this), then in certain cases an integrating factor λ can be found, such that

$$\begin{aligned}\lambda\,\mathrm{d}'w = {} & \mathrm{d}\eta = \lambda A_1\,(x_1, x_2, \ldots, x_n)\;\mathrm{d}x_1 \\ & + \lambda A_2\,(x_1, x_2, \ldots, x_n)\;\mathrm{d}x_2 \;+\; \ldots \\ & + \lambda A_n\,(x_1, x_2, \ldots, x_n)\;\mathrm{d}x_n, \end{aligned} \tag{4}$$

where $\mathrm{d}\eta$ is a total differential.
For two variables, a Pfaffian expression always has an integrating factor. However, if the equation has three or more variables, the existence of an integrating

factor cannot be generally assumed. According to Carathéodory, the condition of integrability can be stated topologically. There is an integrating factor, if and only if there are points in every neighborhood of a given point, which cannot be reached along curves representing solutions of the Pfaffian expression. The connection between this mathematical theorem and his statement concerning adiabatic processes is apparent. Nevertheless, the proof for the global existence of the integrating factor is missing. Carathéodory chose a topological way to prove the existence of an integrating factor for heat, whereas Gyula Farkas showed that these conditions are satisfied for heat in thermodynamics by more physical arguments.

4 The Farkas proof

Gyula Farkas made his considerations upon reading Voigt's book and published a paper in answer to it, which was translated into German as well [10, 11].

4.1 The general θ-function

Thermodynamic systems can be described by a number of state variables $a, b, c, \ldots$ and an *empirical* temperature θ, so the heat flow can be expressed as

$$\mathrm{d}'Q = \Theta\,\mathrm{d}\theta + A\,\mathrm{d}a + B\,\mathrm{d}b + C\,\mathrm{d}c + \ldots, \tag{5}$$

where $\Theta, A, B, C, \ldots$ are Lipshitz-continuous and neither infinite (unphysical due to the limited energy in universe) nor zero (Θ is obviously the heat capacity, as the advanced reader recognizes), and d' indicating that the differential is not total, i.e. its integral is path dependent. As a direct consequence of the Second Law, Farkas proved his lemma. For reversible adiabatic processes (since $\mathrm{d}'Q = 0$), this can be rewritten to get a new Pfaffian expression for θ:

$$\mathrm{d}\theta = -\frac{A}{\Theta}\,\mathrm{d}a - \frac{B}{\Theta}\,\mathrm{d}b - \frac{C}{\Theta}\,\mathrm{d}c + \ldots\ . \tag{6}$$

Theorem 1. *In the adiabatic equation, $d\theta$ is a total differential.*

Proof. First, the *Farkas lemma* is introduced:

Farkas Lemma of Thermodynamics: In a reversible process, no system can go adiabatically into a state to which it can go by only abstracting heat, i.e. by only supplying or removing heat.

Proof. If this were possible, $\theta_i, a, b, c, \ldots$ to $\theta_f, a, b, c, \ldots$ would be an adiabatic process and we could construct a circular process, which only transfers heat from a hot to a cold reservoir. So the lemma holds.

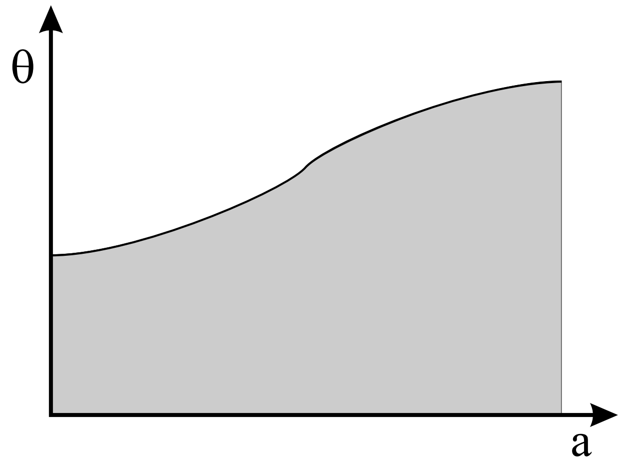


Fig. 1: θ as a function of a. With an a adiabatic process, $\mathrm{d}'Q = 0$, the hatched area cannot be reached due to the *Farkas Lemma of Thermodynamics*.

Farkas Corollary: The temperature θ is always entirely defined by the values of the other state variables. The *Farkas corollary* ensures that the adiabatic equation (6) is integrable, that is

$$\theta(a, b, \ldots) = \theta_0 - \int_{a_0, b_0, \ldots \theta_0}^{a, b, \ldots} \left(\frac{A}{\Theta} \, \mathrm{d}a - \frac{B}{\Theta} \, \mathrm{d}b + \ldots \right). \tag{7}$$

If θ_0 is constant, this is the equation for an adiabatic process. It means that adiabatic processes can be identified by θ_0. Further, fixing the state variables as $a_0, b_0, c_0, \ldots$, and considering them as parameters, θ_0 can be a new state variable. The independent parameter set will be $(a, b, c, \ldots, \theta_0)$, and then the empirical temperature θ, as dependent variable can be written as:

$$\theta = \theta \left(a, b, c, \ldots, \theta_0 \right). \tag{8}$$

Theorem 2. *The θ-function is differentiable by its parameters $a, b, c, \ldots, \theta_0$*

Proof. Expanding the differential of θ by its parameters,

$$\mathrm{d}\theta = \frac{\partial \theta}{\partial a} \, \mathrm{d}a + \frac{\partial \theta}{\partial b} \, \mathrm{d}b + \frac{\partial \theta}{\partial c} \, \mathrm{d}c + \ldots \; . \tag{9}$$

and comparing it with the earlier expression (6), we get for $\theta_0 = const$

$$\frac{\partial \theta}{\partial a} = -\frac{A}{\Theta}, \qquad \frac{\partial \theta}{\partial b} = -\frac{B}{\Theta}, \qquad \frac{\partial \theta}{\partial c} = -\frac{C}{\Theta}, \qquad \ldots \, , \tag{10}$$

and hence for $\theta_0 = const$ all the differentials exist and $\mathrm{d}\theta$ is differentiable.

Lemma 4. *θ is a differentiable function of θ_0 in every point except a set of measure zero.*

Proof. The theorem of Lebesgue on the differentiation of monotonous functions implies the existence of a differential in every point except a set of measure zero. The $\theta = \theta_0$ function is monotonous because of the Second Law. Thus θ is differentiable by θ_0 everywhere except on a set of measure zero, too.

4.2 Constructing the Integrating Factor

To construct the integrating factor, we have to regard *quasi-static adiabatic processes*; quasi-static is used in the sense of Uffink (see [8]), meaning that changes in the process should take place so slow that the system can be regarded in equilibrium up to a negligible error. Therefore, we have to take θ_0 into our considerations as well. So the expansion of the θ-differential is

$$\mathrm{d}\theta = \frac{\partial\theta}{\partial a}\,\mathrm{d}a + \frac{\partial\theta}{\partial b}\,\mathrm{d}b + \frac{\partial\theta}{\partial c}\,\mathrm{d}c + \ldots + \frac{\partial\theta}{\partial\theta_0}\,\mathrm{d}\theta_0\ . \tag{11}$$

With this, we can rewrite 5:

$$\begin{aligned}\mathrm{d}'Q = \Theta&\left(-\frac{A}{\Theta}\,\mathrm{d}a - \frac{B}{\Theta}\,\mathrm{d}b - \frac{C}{\Theta}\,\mathrm{d}c - \ldots + \frac{\partial\theta}{\partial\theta_0}\,\mathrm{d}\theta_0\right)\\ &+A\,\mathrm{d}a + B\,\mathrm{d}b + C\,\mathrm{d}c + \ldots &(12)\\ =\ &\Theta\frac{\partial\theta}{\partial\theta_0}\,\mathrm{d}\theta_0 &(13)\end{aligned}$$

In each interval, where $\partial\theta_0/\partial\theta$ does exist (according to the Lebesgue-theorem),

$$\mathrm{d}'Q = \Theta\frac{\partial\theta}{\partial\theta_0}\,\mathrm{d}\theta_0, \tag{14}$$

In this form

$$\Theta\frac{\partial\theta}{\partial\theta_0} \tag{15}$$

is the integrating factor. This approach shows as well that the integrating factor exists almost everywhere, but it is not unique: Since the empirical temperature scale θ can be chosen freely, there are many empirical entropies s as well as many integrating factors, although we have the same physics.

We can introduce new variables $s = f(\theta_0)$ (called empirical entropy), and φ is the new the integrating factor:

$$\varphi = \Theta\frac{\partial\theta}{\partial\theta_0}\frac{\mathrm{d}\theta_0}{\mathrm{d}s}. \tag{16}$$

Now we can rewrite the equation to its final and well known form:

$$\mathrm{d}'Q = \varphi\,\mathrm{d}s \tag{17}$$

4.3 Absolute Temperature and Entropy

Theorem 3. *There is a unique (absolute) temperature, which yields additive entropy. It is called the absolute temperature scale.*

Proof. Consider two bodies in thermal equilibrium. The heat flow can be expressed as follows:

$$\begin{aligned} \mathrm{d}'Q_1 &= \Theta_1 \,\mathrm{d}\theta + A_1 \,\mathrm{d}a_1 + B_1 \,\mathrm{d}b_1 + C \,\mathrm{d}c_1 = \varphi_1 \,\mathrm{d}s_1 \\ \mathrm{d}'Q_2 &= \Theta_2 \,\mathrm{d}\theta + A_2 \,\mathrm{d}a_2 + B_2 \,\mathrm{d}b_2 + C \,\mathrm{d}c_2 = \varphi_2 \,\mathrm{d}s_2 . \end{aligned}$$

Here the Zeroth Law was assumed, that is there are no long-range interactions. The two bodies are described only by their own state variables, i.e. they have to be independent systems. The heat flow for the united system is the sum of the flow for each of the systems:

$$\mathrm{d}'Q = \mathrm{d}'Q_1 + \mathrm{d}'Q_2 = \varphi_1 \,\mathrm{d}s_1 + \varphi_2 \,\mathrm{d}s_2 = \Phi \,\mathrm{d}S. \tag{18}$$

Substituting a_i by s_i and expanding $\mathrm{d}S$ (which again has to be continuous by the requirements), we get:

$$\mathrm{d}S = \frac{\partial S}{\partial s_1}\mathrm{d}s_1 + \frac{\partial S}{\partial s_2}\mathrm{d}s_2 + \frac{\partial S}{\partial \theta}\mathrm{d}\theta + \frac{\partial S}{\partial b_1}\mathrm{d}b_1 + \frac{\partial S}{\partial b_2}\mathrm{d}b_2 + \dots . \tag{19}$$

Comparing with 18 and considering that $s_1,\ s_2,\ \theta,\ b_1,\ b_2,\ \dots$ are independent, we get:

$$\frac{\partial S}{\partial \theta} = 0, \quad \frac{\partial S}{\partial b_1} = 0, \quad \frac{\partial S}{\partial b_2} = 0, \quad \dots, \tag{20}$$

and hence S depends only on s_1 and s_2. φ_1/Φ and φ_2/Φ are functions of s_1 and s_2 only, since

$$\mathrm{d}S = \frac{\varphi_1}{\Phi}\mathrm{d}s_1 + \frac{\varphi_2}{\Phi}\mathrm{d}s_2 \tag{21}$$

has to depend upon s_1 and s_2 only. The ratios φ_1/Φ and φ_2/Φ, too, have to be functions of s_1 and s_2 only, e.g.:

$$\frac{\varphi_2}{\Phi} = \frac{\varphi(\theta, s_1)}{\Phi(\theta, s_1, s_2)} \tag{22}$$

In order that this ratio is dependent only of s_1 and s_2, the θ-dependence has to factor out in a *universal* function:

$$\begin{aligned} \varphi_1 &= f(\theta)\Psi(s_1) \\ \varphi_2 &= f(\theta)\Psi(s_2). \end{aligned} \tag{23}$$

This way we can define a new S in the form

$$\begin{aligned} \mathrm{d}S_1 &= \Psi(s_1)\,\mathrm{d}s_1 \\ \mathrm{d}S_2 &= \Psi(s_2)\,\mathrm{d}s_2, \end{aligned} \tag{24}$$

and we gained a new integrating factor, which is a *universal* function of the empirical temperature for both systems:

$$\varphi_1 = \varphi_2 = T(\theta). \tag{25}$$

This integrating factor is a universal temperature scale defined for all systems (and hence the entropy has a universal scale as well) with the property

$$\mathrm{d}'Q = T\,\mathrm{d}S. \tag{26}$$

Now we can have a closer look to the set where θ as a function of θ_0 is not differentiable (the set of measure zero mentioned beforehand). For the heat we have now:

$$\mathrm{d}'Q = T(\theta)\,\mathrm{d}S(\theta, a, b, c, \ldots) \tag{27}$$

$T(\theta)$ is a universal function. But $S(\theta, a, b, c, \ldots)$ may change its slope at certain points, since $\mathrm{d}S$ does not have to exist everywhere. From physics, we know that this is indeed a property of entropy: It is not differentiable on phase transitions. So we can identify these points with the phase transitions, and we see that the Farkas theory reflects a further piece of physics in mathematical facts!
Thus Farkas' construction leads directly from Clasius' or from Kelvin's postulate and the Farkas lemma to the mathematically rigorous proof of the existence of an integrating factor and its identification with absolute temperature, and the definition of a entropy function.

5 Examples

5.1 The Ideal Gas

To illustrate this proof, we provide first an easy example how an integrating factor can be found in a concrete problem, the ideal gas. For convenience, we will consider one mole of an ideal gas with T (ideal gas temperature) being the empirical temperature θ and a (the first state variable) being the volume V. $\Theta = c_v = 1.5R$ and R is the universal gas constant. Further, $a = p$, where $p = RTV^{-1}$ is the pressure.
We have to start with the expression for the heat:

$$\mathrm{d}'Q = c_v\,\mathrm{d}T + \frac{RT}{V}\,\mathrm{d}V. \tag{28}$$

The adiabatic equation is

$$\frac{\mathrm{d}T}{T} = -\frac{R}{c_v V}\,\mathrm{d}V. \tag{29}$$

Integrating this yields the equation of adiabatic curves

$$\ln T = \ln V^{-\frac{R}{c_v}}, \tag{30}$$

which is equivalent with

$$TV^{\frac{R}{c_v}} = T_0 V_0^{\frac{R}{c_v}}. \tag{31}$$

So we see that T_0 is the constant of the adiabatic surface, which we labeled θ_0 before: it gives the temperature value belonging to V_0. From now on, V_0 is a fixed parameter, and we can solve for T_0:

$$T = T_0 \,(V/V_0)^{-\frac{R}{c_v}}. \tag{32}$$

To get the integrating factor, we proceed like in 16:

$$\frac{\partial \theta}{\partial \theta_0} = \frac{\partial T}{\partial T_0} = (V/V_0)^{-\frac{R}{c_v}} \tag{33}$$

and thus the

$$d'Q = c_v T \frac{\mathrm{d}T_0}{T_0} \tag{34}$$

Now we can calculate the entropy function, $d'Q = T\,\mathrm{d}S$:

$$S = \int \frac{c_v}{T_0}\,\mathrm{d}T_0 = c_v \ln(TV/V_0)^{\frac{R}{c_v}} = R \ln T^{\frac{c_v}{R}} V. \tag{35}$$

So we have shown that the ideal gas temperature is the absolute temperature, and we found a relation, which expresses the absolute entropy in terms of the function V, T for the adiabatic surfaces.

5.2 The Radiation

An example, which is somewhat less trivial, so that it suits better as demonstration, is measuring the temperature of a gas by its heat radiation (from which we know, by the Stefan-Boltzmann law, that the radiation goes with the fourth power of the temperature. The energy is $U = \sigma\, T^4 V$. Now we select $\theta = \sigma\, T^4$)as empirical temperature and the volume V the first variable a. We will show that the calculations will lead to the real temperature form.
Again, we apply the Poisson formula for adiabatic processes (but this time the radiation temperature is the empirical temperature scale):

$$\mathrm{d}'Q = c_v\,\mathrm{d}\theta + 4/3\theta dV \tag{36}$$

where

$$c_v = V \tag{37}$$

Regrouping and integration yield the equation of adiabatic curves

$$\theta = \theta_0 (V/V_0)^{-4/3}, \tag{38}$$

This equation describes the adiabatic surfaces. Now we can calculate the integrating factor:

$$d'Q = V(V/V_0)^{-4/3}\,\mathrm{d}(\theta(V/V_0)^{4/3}, \tag{39}$$

that is

$$d'Q = \frac{4}{3}\theta^{1/4}\,\mathrm{d}(\theta^{3/4}V) \tag{40}$$

and by the requirement, that the universal temperature has to factor out, we see that this must be

$$T = \left(\frac{\theta}{\sigma}\right)^{1/4} \tag{41}$$

and the entropy is,

$$S = \frac{4}{3}\sigma^{1/4}\theta^{3/4}V = \frac{4}{3}\sigma^{1/4}U^{3/4}V^{1/4} \tag{42}$$

As expected, the absolute temperature is proportional to the fourth root of θ, we got back the well-known expressions for absolute entropy and absolute temperature of the radiation.

References

[1] Farkas, G., Mathematikai és Természettudományi Értesítő **16**, 361 (1896).
[2] Farkas, J., Journal für die Reine und Angewandte Mathematik, **124**, 1 (1901).
[3] Clausius, R., *Abhandlungen über die mechanische Wärmelehre* (Vieweg und Sohn, Braunschweig 1864).
[4] Zeuner, G., *Grundzüge der mechanischen Wärmetheorie* (1866), 2nd Ed.
[5] Voigt, W., *Kompendium der Theoretischen Physik* (Verlag von Veit und Comp. 1895).
[6] Caratheéodory, C., Math. Ann. **67**, 366 (1909).
[7] Bernstein, B., J. Math. Phys. **1**, 222 (1960).
[8] Uffink, J., in *Entropy*, edited by A. Greven, G. Keller, and G. Warnecke (Princeton University Press, 2001), pp. 121-147.
[9] Martinás, K., Brodszky, I., Periodica Polytechnica Ser. Chem. Eng. **44**, 17 (2000).
[10] Farkas, G., Mathematikai és Phisikai Lapok, **IV**, 7 (1895).
[11] Farkas, J., Mathematische und Naturwissenschaftliche Berichte aus Ungarn **XII**, 282 (1895)

History and Education

Gyula Farkas and the Franz Joseph University from Cluj/Kolozsvár

György GAAL

RO 400015 Cluj–Napoca
Str. Republicii 43
Romania

Abstract. This study in its first chapter presents the short history of higher education in Cluj/Kolozsvár. The first attempt dates back to 1581, the second one to 1774-1775. Only the third foundation that from 1872 created a long-lasting institute. This is the Franz Joseph University, which had four faculties, one being that of Mathematical-Natural Sciences. The beginning of this Faculty and its professors of Mathematics and Physics are presented in the second chapter. One of the most outstanding professors was Gyula Farkas who taught Theoretical Physics between 1887–1915. His activitiy as member of teaching staff is described in the third chapter. It is mentioned his role as a dean and as rector of the university. Most of the dates are taken from year-books of the university. The last chapter charecterizes Gyula Farkas as a lecturer and educator, gives a list of his multiplied courses, enumerates his best students. In the Appendix the title of all courses held by professor Farkas at Cluj/Kolozsvár university are to be found.

AMS 2000 subject classifications: 01A70
Key words and phrases: Gyula Farkas, Franz Joseph Univesity, higher education

1 Higher Education in Cluj/Kolozsvár

The beginnings of higher education in Cluj/Kolozsvár dates back to the end of the 16th century[1]. Then István Báthory, prince of Transylvania and king of Poland decided to strengthen Catholicism in Transylvania, because most of the country became protestant at that time. So in 1579 he sent twelve Jesuits to Transylvania and ordered them to open a college in the neighbourhood of Cluj/Kolozsvár, at the former Benedictine Abbey of Kolozsmonostor. In 1580 this college was moved inside the town into the Farkas Street. The wholly protestant town protested against this decision. On the 12th May 1581 Báthory signed a founding charter according to which he raised the college to the rank of university. It could promote students to baccalaureat's, magister's and doctor's degree just similar to any foreign academy. Two years later the prince founded

[1] Referring to the history of higher education in Transylvania and the history of Franz Joseph University see: *A kolozsvári M. K. Ferencz-József Tudomány-Egyetem.* In: A felsőktatásügy Magyarországon. Budapest, 1896. pp. 247–405.; Márki Sándor: *A M. Kir. Ferencz József Tudományegyetem története 1872–1922.* Szeged, 1922.; Gaal György: *Egyetem a Farkas utcában.* A kolozsvári Ferenc József Tudományegyetem előzményei, korszakai és vonzatai. Kolozsvár, 2001.

a students' hostel, called seminary, housing 150 persons. For this a new three story building was raised. In the golden age of this academy about 20–25 professors were teaching more than 300 students. They were applying the famous educational system and curriculum of Jesuits. The protestant population was offended by the missionary activity of the Jesuits, so they were driven out from the city first in 1603, then in 1605 their buildings being destroyed.

The second attempt to create a university at Cluj/Kolozsvár is linked to the name of empress Maria-Theresia. Since 1690 Transylvania belonged to the Austrian Empire, the Austrian rulers were in the same time princes of Transylvania. Vienna supported Catholicism and disliked the habit of protestant churches to send their best students to German, Dutch or even British academies from where they often returned with "dangerous" ideas. So when the Jesuit Order was prohibited by the Pope in 1773 the empress decided to found a university in their buildings. Though the Protestants were invited to collaborate they were rather suspicious, because they knew: the leading offices will be given to Catholics and probably their students will get no passports any more to foreign universities. Even so the empress founded a Faculty of Law in 1774 and a Faculty of Medicine in 1775 where Protestants could get teaching jobs. Meanwhile the Faculties of Philosophy and Divine Studies remained purely of catholic spirituality their teachers being Piarist fathers. This new institute was called Royal Academic Lyceum but mostly they mentioned it as a university. Its Faculty of Law functioned till the 1848 revolution. Then after a long brake it was reestablished as a Royal Academy of Law in 1863. The Faculty of Medicine in 1817 became a separate Medical-Surgical Institute.

Following the 1867 compromise between Austria and Hungary as well the integration of Transylvania into Hungary there was need to create a second Hungarian university. At the beginning Bratislava/Pozsony and Cluj/Kolozsvár were in competition for this institute. But the second town was in a much more favourable position. Two faculties (Law, Medicine) were already existing. The former Transylvanian government buildings became free and were more or less suitable for a university. Then the Transylvanian Museum Society (founded in 1859) had rich scientific collections, a large library which were at the disposal of the students.

The Bill of Foundation was brought into the Hungarian Parliament in April 1870 by József Eötvös, the famous novelist (the father of the physicist Loránd Eötvös) then Minister of Education. After two years of waiting the next minister, Tivadar Pauler asked the permission of the emperor/king Franz Joseph for opening the Cluj/Kolozsvár University. This was given on the 29th of May 1872. So the preparations could begin. The Bill was discussed in the Hungarian Parliament only in the autumn of that year. The newly founded university had four faculties (Law, Medicine, Philosophy, Mathematical-Natural Sciences) and 42 teaching chairs. There were about 120 applications for the chairs. The opening ceremony and the installation of the first rector took place on the 10th of November 1872. It is worth mentioning that this new university had a separate faculty for natural sciences. The Hungarians mostly followed the German school-system, where – except the Tübingen University – at that time Natural Sciences were integrated into the Faculty of Philosophy. So it happened at Budapest University, too. The Cluj/Kolozsvár Faculty of Mathematical-Natural Sciences had eight teaching chairs. In autumn 1872 only seven were occupied: Elementary Mathematics – Sámuel Brassai; Higher Mathematics – Lajos Martin; Experimental Physics – Antal Abt; Geology – Antal Koch; Chemistry – Antal Fleischer; Botanics – Ágost Kanitz; Zoology – Géza Entz. For the eighth chair, that of Mathematical or Theoretical Physics no candidate of corresponding studies was found. The most venerable member of the whole teaching staff was Sámuel Brassai, a polyhistor, expert in about ten sciences, who was immediately elected as prorector. He was a full member of the Hungarian Academy of Science. Martin was corresponding member of the Academy.

2 Mathematics and Physics departments and teachers

The vacant chair for Theoretical Physics was first occupied in 1874 by Mór Réthy, a young man of science who just completed his studies at German universities. At the beginning he was extraordinary, from 1876 full professor. He was a great scientist, at the age of 30 in 1878 he was already elected corresponding member of the Academy. He also studied the heritage of the two Bolyais. In 1884 he preferred to be transferred to the chair of Elementary Mathematics which became vacant after the retirement of Brassai. Two years later he was invited to the Technical University from Budapest. The department for Theoretical Physics in 1884 was occupied by Gyula Vályi, one of the first graduates of the Cluj/Kolozsvár University, its privatdocent at that time. He worked there for two years as extraordinary professor. Then he also preferred the Elementary Mathematics, so he was promoted there in January 1887 as full professor. As a result there was again vacancy at Theoretical Physics Department, but this time there appeared a candidate who, though a great scientist, remained for nearly three decades devoted to this chair and this university: Gyula Farkas. From 1881 he was a privatdocent of Budapest University, on the 8th January 1887 he occupied the chair as extraordinary professor, on the 23rd of March 1888 he was promoted full professor.

At that time there were still eight departments at the Faculty, but some of the professors were already belonging to the second generation. At Elementary Mathematics professor Vályi was teaching, a great mathematician, a commentator of Bolyai's Appendix. In 1891 he entered the Academy as a corresponding member. At Higher Mathematics still Martin was the professor, he became famous for his experiments in aeronautics. After his death, in 1897 the privatdocent Lajos Schlesinger was promoted professor: a mathematician with German studies and international fame. In 1902 he was elected corresponding member of

Academy. He is considered the founder of the scientific library of the mathematical departments. The professor of Experimental Physics was Antal Abt since the foundation. Not only a scientist but a good organizer and a popular lecturer. He succeeded to create an Institute of Physics with all the necessary experimental instruments. As well he put the bases of the Cluj/Kolozsvár school of Physics. Just before his death (1902) he could install his institute into the new central building of the university, where on the second floor there were about 13 rooms on his disposal, on the underground level some workshops were helping the experimental activity. The next professor and head of the institute was Károly Tangl, a former assistant of Loránd Eötvös, privatdocent of Budapest University. Also a good methodologist and a man of science, who in 1908 was elected corresponding member of the Academy. At this department at the very beginning there was an assistant. At the end of the century Péter Pfeiffer helped the professor, later he became first assistant, then privatdocent (1902), and beginning with 1904 he founded the third department in this discipline, that of Practical Physics.[2]

In the school-year 1886/1887 when Gyula Farkas entered the teaching staff there were four privatdocents at the faculty, none of Physics, and six assistants, from which Pfeiffer was a physicist. The faculty was lead by the dean: Antal Abt. His deputy being Lajos Martin. At that time, according to the law each rector or dean next year served as a deputy. The Faculty of Mathematical-Natural Sciences was the smallest, with few teachers so they were quite often elected deans. The students' number was also small. From 456 only 30 were studying at this faculty. While in the next decade the most popular disciplines as Chemistry and Experimental Physics were taken up yearly by more then 50 students, Theoretical Physics had 12–15 students, being considered the most difficult subject.

3 Gyula Farkas as member of the teaching staff

The Franz Joseph University from Cluj/Kolozsvár published its year-books in three volumes: Acta (speeches and documents); Almanac (not a calendar, but the presentation of the Senate, enumeration of all the teachers with their titles and the institutes belonging to the university); Timetable (separately for each semester, containing all the delivered courses and seminars, the teacher's name, the classroom as well as the home-address of the teachers).[3] From these volumes

[2] For short biographies see: *Százhuszonöt éve nyílt meg a kolozsvári tudományegyetem.* Emlékkönyv. Vol. I–II. Piliscsaba, 1997.

[3] Acta Reg. Scient. Universitatis Claudiopolitanae Francisco-Iosephinae Anni MDCCCLXXXVII–LXXXVIII. (1887–1918); A kolozsvári Ferencz-József Tudományegyetem Almanachja MDCCCLXXXVI–LXXXVII-ről. (1886–1916); A kolozsvári Ferencz-József Tudományegyetem Tanrendje az MDCCCLXXXVII–LXXXVIII. tanév I-ső felére. (1887–1915). Between 1888 and 1910 Almanach and Tanrend were published in a single volume.

we can learn a lot about Gyula Farkas' activity. It is easy to follow how his appreciation grows, how he becomes more and more honorable, getting functions and titles, then how he withdraws more and more from university life, addicting himself only to science and teaching.

When Farkas arrived to Cluj/Kolozsvár university in 1887 he got the rank of extraordinary professor. This meant that his scientific activity was not yet worthy enough for his job. Otherwise he had all the duties of a full professor. Maybe his salary was smaller. All the extraordinary teachers after one or two years' activity were promoted, so it happened with Farkas. He was immediately appointed as member of two institutes/committees connected with his faculty. The university in those times prepared scientists. Those who wanted to become schoolteachers had to listen some courses at the Teacher-training Institute, where 15 professors from the faculties of Philosophy and Mathematics delivered some lectures. Those who studied some years at this institute had to be examined by the Schoolteachers' Examining Committee to get a diploma. During his whole university career Farkas was member of these two staffs. From 1898 till 1908 he was vice-president of the Examining Committee. Though he did not deliver – except the very beginning of his activity – training lessons. This kind of activity usually was done by the professor of Experimental Physics.

The "Department of Theoretical Physics" during this period meant one single chair of a professor. He had no assistants or other substitute teachers. He could propose the promotion of privatdocents[4] in his field, but only Lipót Fejér obtained this degree. It is interesting to see, that the Department of Theoretical Physics is mostly connected to the two departments of mathematics not to that of Experimental Physics. Farkas usually delivered 7–9 classes a week, from which 2 were seminars. Only in the 1901/1902 school–year was founded a Seminar in Mathematics. It had three directors (Vályi, Farkas, Schlesinger), Schlesinger being its real manager. This meant a room with a library of special issues. Beginning with 1903 a 4th year student, called tutor took care of this Seminar. In 1905 Lipót Fejér became the tutor and the common assistant of the three professors. Next year he was already promoted privatdocent, then in 1911 he became extraordinary professor of Higher Mathematics, but after some months he left for Budapest. No other person is mentioned as tutor at this Seminar.

Like other professors, Gyula Farkas was periodically elected dean of his faculty. It seems that he was a very accurate office-holder and could manage well the small faculty. This is why he was seven times elected dean: 1889/90, 1892/93, 1893/94, 1896/97, 1897/98, 1898/99, 1902/903. Three times he was repeatedly

[4] Privatdocents, according to German university-system, were promoted at the proposal of two ordinary professors appreciating their whole scientific activity, by the decision of the faculty's council. Privatdocents were not employees of the university, but they could deliver some facultative courses. In case of vacancy the privatdocents had a good chance to be appointed.

elected, probably no one wanted to overtake the office. According to the internal law each dean next school-year became deputy-dean. But because of reelections and other special cases this law could not always be respected. In such way Farkas was deputy-dean in five school-years: 1890/91, 1894/95, 1899/900, 1901/902, 1903/904.

As dean he had four times very honourable tasks representing not only his faculty but the whole university. In December 1892 the University of Padua feasted the 300th anniversary of Galileo Galilei's introduction as a professor. Farkas presented the greetings of the Cluj/Kolozsvár University.[5] There he got the title of doctor honoris causa. On the 24th of June 1897 Sámuel Brassai the retired professor, former rector of the university died. Not only the whole town and university was mourning the great scientist but the whole country. At the funeral Farkas delivered a speech. In the same year, on 6th July the university promoted doctor honoris causa archduke József on the proposal of Mathematical Faculty. So it became Farkas' job to handle to the member of the royal family the diploma delivering a short speech. In January 1903 the Franz Joseph University organized a festivity commemorating the centenary of János Bolyai's birth. At this occasion a memorial plaque was put on his birth-house. Farkas was one of the organizers, and he was the speaker of the unveiling. There is a fifth occasion, too, where Farkas represented not so much his faculty but the Academy of Science: it was the reburial of János Bolyai on the side of his father in Marosvásárhely Calvinist graveyard in July 1911. Farkas was competent in both occasions because he wrote a study on Bolyai's theory.

The summit of his university career can be considered his election as a rector for the 1907/908 school-year. Thus for one year he became the president of the Senate, its members being the prorector, the four deans and the deputy-deans. He had to preside at the meetings of the Senate and at festival occasions. He had to deliver an inaugural address, then next September to give an account of his activity. Finally he had to arrange all his speeches and other documents of the school-year into a volume (Acta) of the year-books. We may state that during his rectorship no special events happened.

Gyula Farkas delivered his inaugural speech on the 22nd of September 1907 in the aula of the university at the opening ceremony of the school-year.[6] This is a very long essay of 30 pages, about the development of Theoretical Physics in the last two decades, since he has been appointed professor of this subject. In the same time expresses his creed about the tasks of university teachers and students. Before enumerating the new discoveries, theories in each field of his discipline he points out: there was such a quick progress and huge development

[5] Farkas Gyula: A Galilei-ünnep Páduában. Természettudományi Közlöny, vol. XXV. (1893) pp. 196-201.

[6] Published in Acta 1907-1908 and also as an extract: *Beszéd, mellyel Farkas Gyula [...] a kolozsvári Magyar Királyi Ferencz József Tudományegyetem e[z]i[dei] rektora az 1907/8. tanévet megnyitotta.* Kolozsvár, 1907.

that he had to study all the time, to understand the novelties, and then repeatedly rewrite his courses. In the end he confesses: he can't understand those students who after graduating their studies, consider that they had acquired all the necessary knowledge. The university just enables its students to study and to understand the problems of a discipline. They must continue to study all their life, they must deepen their knowledge. Another occasion to deliver a speech was on the 29th of May 1908, which was the anniversary of the foundation.[7] Each year on this day the best competition essays were awarded. The winners of some scholarships were announced. Rector Farkas in his opening speech mentioned that it is the 35th occasion to give the awards and showed what are the funds which grant the necessary sums.

His giving account speech was delivered on the 27th of September 1908, before the opening address of the new rector.[8] It is a very well articulated text with 18 sub-titles and 41 appendices. These are mostly documents and statistics. We just point out that in the first semester there were 47 ordinary, 5 extraordinary professors, 43 privatdocents in the teaching staff, from the 2404 students only 164 were listening courses at the Faculty of Mathematical and Natural Sciences. Two important new buildings were nearly completed: that of the Zoological Institute (Mikó-Garden) and of the University Library. Both of them were inaugurated with the participation of the Minister of Education on the 18-19th May 1909, when Farkas was the prorector of the university. It is worth to mention that in the period 1898-1908 he was member of the library board.

It is often mentioned[9] that Gyula Farkas was attracting the best mathematicians to the Cluj/Kolozsvár university. He even wrote letters inviting them. So during the last five years of his activity his colleagues were Lajos Schlesinger, Lipót Fejér, Alfred Haar, Frigyes Riesz who later on founded the world-famous Hungarian School for Mathematics. Their common Seminar was probably the birthplace of this scientific school. They all were members of the famous Circolo Matematico from Palermo.

Most of the professors in those times retired at the age of 70. Gyula Farkas suffered from eye-disease. He considered that he can't carry out his duties

[7] Acta 1907/8. Fasciculus II. *Beszédek, amelyek a kolozsvári M. K. Ferencz József Tudományegyetem alapítása XXXVI. évfordulójának ünnepén 1908. május hó 29-én tartattak [...].* Kolozsvár, 1908.

[8] Acta 1908/9. Fasciculus I. *Beszédek, amelyek a kolozsvári M. K. Ferencz József Tudományegyetem 1908/9. tanévi rectora és tanácsa beiktatása és a tanév megnyitása alkalmából 1908. évi szeptember hó 27-én tartattak.* Kolozsvár, 1908. 1-144. (Rectori beszámoló beszéd. Mondotta [...] Farkas Gyula mint az egyetem kormányáról lelépő 1907/8. tanévi rector 1-22 + Mellékletek 22-144.)

[9] E.g. Gábos Zoltán: "*A természet a matematika nyelvén szól hozzánk*" Természet Világa, Vol. 128. (1997) pp. 290-293.; Prékopa András: *Farkas Gyula élete és munkásságának jelentősége az optimalizálás elméletében.* In: Farkas Gyula élete és munkássága. Budapest, 2003. pp. 9–26.

properly. Therefore he resigned at the age of 68, after 28 years of teaching at Cluj/Kolozsvár. During his last school-year being mostly on sick-leave professors Tangl and Haar held his courses. The decision on his retirement is dated on the 23rd of October 1915. The rector of 1915/16 school-year, professor Károly Tangl commemorates him with these words: "During the first semester dr. Gyula Farkas, professor of Mathematical Physics at our university for 28 years has retired. It was not easy to part with this beloved teacher of our university, who during his long activity won our appreciation and respect. His courses will remain an example for accuracy and clearness. He was able to make clear the most difficult problems, he avoided any extra word, so his lectures were classical examples for conciseness and preciseness. His valuable studies in Mechanics and Electrodynamics will remain a source of esteem for our university. We highly esteemed his love of fair play which is reflected in all his deeds, and was especially precious in managing our university. He took leave of the university, but it may give us consolation that he continues to work for Hungarian science, would God let him for many years!"[10] At his retirement Farkas was awarded with the middle-cross of the Franz Joseph-order.

4 Gyula Farkas as lecturer and educator

From the official Timetable of the university we know the title of all his courses. From these titles a physicist could probably draw some conclusions upon the interests of professor Farkas. It was up to him to decide the number of weekly classes between 7 and 10, as well the number of courses and seminars. In the first school-years he proposed two courses (3+2 classes) and a seminar (2 classes) and sometimes a public free lesson (1 class) which was ment for all the students. Beginning with the 1890ies he delivered a basic course of 5 classes (each working day at noon) and had two practical classes connected with the course (usually on Wednesday afternoon). Sometimes in one of the semesters he had a second course of 1 or 2 classes for beginners. In this period all his lessons were held in the old Jesuit building at classroom no. X. In the 20th century in the new building he always delivered his lessons and seminars in the 2nd classroom for mathematics (second floor, north wing, just opposite to the main stairs). From this time up to the end of his activity each semester he had two courses: one of 4 classes for advanced students, one of 3 classes for beginners and a seminar of 2 classes without a special title (Seminar in Mathematical Physics). His courses were always held at noon. The Study of Vectors and Dynamics were his two basic courses which nearly each year appear in the timetable.

One of his best pupils, his successor Rudolf Ortvay in his commemorating speech delivered in the meeting of the Hungarian Academy of Science in 1932 characterizes in this way the lecturer Farkas: "I have to appreciate his courses,

[10] Acta 1916/17. Fasciculus I. p. 19.

which he carefully prepared for lithography. His courses comprehended the whole area of Theoretical Physics, but their scheme was different from the usual one. He put a stress on the accurate and correct explanation of the fundamental concepts. He discussed only such phenomena which were already settled enough for teaching. His style was so concise that not a word could be omitted. Therefore these texts are not easy to understand and less suitable for beginners as an introduction into the science. But those whose mind is already cultivated and who appreciate the general view and the rigour of science as well the logical structure of a discipline can learn a lot or even find delight in them. But these features and the lack of any illustration, descriptiveness make easy to understand why these courses were not too popular."[11]

In professor Ortvay's bibliography of Gyula Farkas one printed and five multiplied courses are enumerated. Professor Zoltán Gábos knows about eleven titles, some of them have three editions. He does not refer to his sources but mentions that he has not seen them all. We checked the two great Cluj/Kolozsvár libraries (University, Academy) as well the Library of the Faculty of Mathematics and Computer Science and compared the volumes with the titles mentioned in the yearly bibliography of the teachers published in the Acta-volumes and with the list of Zoltán Gábos.[12] In Acta 1891/92 three multiplied courses are mentioned without any title, Gábos enumerates four courses from 1889-1891. None of them is available in our libraries. In 1901 there was published as a sequence of extracts the single printed course: *Vector-tan és az egyszerű inaequatiok tana. – Theory of vectors and of simple inequations* (Kolozsvár, 1901. pp. 1–165.). In Acta 1907/908 again we can read about the supervision of the text of three courses before lithography. The titles are: *A mechanika alaptanai – Fundamentals of Mechanics* (missing from libraries); *Az energia átalakulásai – Energy transformations* (Kiadják Thomay János, Somogyi István. 1906/7 I. félév. pp.1–114.); *Az energia terjedése – Energy propagation* (Kiadja Bendessy György. Leírta Nagy Imre. 1906/7 II. félév. pp. 1–227). Acta 1908/909 again enumerates two titles: *Különös mechanika – Special Mechanics* - its real title is *Analytikus mechanika – Analytical Mechanics* (1907-8 I. félév. pp. 1–208); *Erőtan – Force theory* (missing). At last in Acta 1914/15 we find six titles, even the number of printed sheets is mentioned: *Vektortan – Vector theory* (24 sheets) - missing; *A mechanika alaptanai – Fundamentals of Mechanics* [27 sheets] (1913-14. tanév II. felében. pp. 1–192.); *Analytikus mechanika – Analytical Mechanics* [27 sheets] (1913-14. tanév I. felében. pp. 1–214.) ; *Erőtan – Force theory* [38 sheets] (1913-14. tanév II. felében. pp. 1–308.); *Az energia átalakulásai – Energy transformations* [33 sheets] (1912-13. I. félév. pp. 1–267.); *Az energia terjedése*

[11] *Farkas Gyula rendes tag emlékezete.* Írta Ortvay Rudolf l[evelező] tag. Felolvasta a Magyar Tudományos Akadémiának 1932. évi december hó 19-én tartott összes ülésén. Budapest, 1933. pp. 31-32.

[12] See: Gábos Zoltán's above quoted article.

– *Energy propagation* (29 sheets) - missing. We may state that all the courses published in the 20th century are available in one of their editions.

We know very little about the students and disciples of Gyula Farkas. Very few students took up his courses, mostly mathematicians and physicists. Only archive records may complete the list of those who were examined by him. The Acta volumes preserved us the names of those students who compiled a prize-winner paper in his subject. Each year each professor could propose some subjects for students. The author of the papers had a code-word, his real name was made public only if he won a prize. In 1888-1890 twice Péter Szabó won prizes, he later obtained the doctor's degree and became professor at Budapest, he was a Bolyai-researcher. In 1893-1895 Pongrácz Kacsóh was a prize-winner. He also took the doctor's degree, but became famous as a composer. In 1895/96 two papers were mentioned in his field, the authors are: Emil Korbuly and Sándor Nagy. A year later György Kaufmann is the prize-winner. After a long brake there is a future physicist who got the prize for Experimental Physics in 1910/11, but the proposal was made by Farkas: he is Zoltán Gyulai who soon would become the assistant of professor Tangl, then ordinary professor of this university. In 1911-1913 twice Vasile Lupan (Lupan Vazul) was awarded for Theoretical Physics papers.

One of his best students was Rudolf Ortvay, who after graduating became assistant (1909-1912) to professor Tangl. Then he left for Germany to specialize himself. He got the privatdocent's degree in 1915 in Theoretical Physics at Cluj/Kolozsvár. Probably one of the proposers was Farkas. In such a way Gyula Farkas left a successor when he retired. On the 16th of August 1916 Rudolf Ortvay was appointed extraordinary professor of Theoretical Physics. Farkas immediately moved to Budapest.[13]

[13] Farkas very often changed his home: 1887 – Felsőszén u. 27., 1888 – Belközép u. 9., 1890 – Belszén u. 13., 1891 – Főtér 13, 2nd floor, 1894 – Belközép u. 22., 1896 – Külmagyar u. 7., 1900 – Főtér 26., 1902 – Szentegyház u. 23., 1904 – Karolina tér 3., 1905 – Sétatér u. 4. 1st floor, 1906 – Szentegyház u. 36., 1908 – Szentegyház u. 31. 2nd floor. His Budapest-address: I. Enyedi u. 11.

5 Appendix
The List of Courses Held by Professor Gyula Farkas

1887/88 – I.

Az erőfüggvények elmélete (potenciál elmélet) és alkalmazása – *Theory of Potentials and application.* 3 h/week.

A mechanika elemei – *Basis of Mechanics.* 2 h/week.

*Tanárképző gyakorlatok – *Practical teacher training.* 2 h/week.

1887/88 – II.

A potenciál elmélet alkalmazása – *Application of the Theory of Potentials.* 3 h/week.

Mechanika elemei – *Basis of Mechanics.* 2 h/week.

*A fizika abszolút mértékei. (Publicum.) – *Absolute measures of Physics.* 1 h/week.

Tanárképző gyakorlatok – *Practical teacher training.* 2 h/week.

1888/9 – I.

Energetika (A hő, elektromosság és vegyfolyamatok mechanikai elmélete) – *Energetics.* 5 h/week.

*Az előadásokkal kapcsolatos gyakorlatok –*Practical works.* 2 h/week.

1888/9 – II.

Az energia terjedési tüneményeinek elmélete (Rugalmassági tünemények, a hang-, fényterjedés, hővezetés, elektromos áramok, elektromos és mágneses indukció elmélete) – *Theory of propagation of energy.* 5 h/week.

Az előadásokkal kapcsolatos gyakorlatok. *Practical works.* 2 h/week.

1889/90 – I.

Analitikus mechanika – *Analytical Mechanics.* 5 h/week.

*Általános függvénytan az alkalmazásban előforduló nevezetesebb függvényalakok ismertetésével – *General Theory of Functions with main function forms in applications.* 2 h/week.

*Az előadásokkal kapcsolatos gyakorlatok – *Practical works.* 2 h/week.

1889/90 – II.

Potenciális elmélet és alkalmazásai – *Application of the Theory of Potentials.* 5 h/week.

*Az előadásokkal kapcsolatos gyakorlatok – *Practical works.* 2 h/week.

1890/91 – I.

Energetika – *Energetics.* 5 h/week.

*Az előadásokkal kapcsolatos gyakorlatok – *Practical works.* 2 h/week.

1890/91 – II.

Az energia terjedési tüneményei – *Propagation of energy.* 5 h/week.

*A fény elektromágneses elmélete. Publikum – *Theory of Electromagnetism.* 1 h/week.

*Az előadásokkal kapcsolatos gyakorlatok *Practical works.* 2 h/week.

1891/92 – I.

Analitikus mechanika – *Analytical mechanics.* 5 h/week.

*Az előadásokkal kapcsolatos gyakorlatok – *Practical works.* 2 h/week.

1891/92 – II.

Potenciális elmélet és alkalmazásai – *Potential Theory and applications.* 5 h/week.

*Az előadásokkal kapcsolatos gyakorlatok – *Practical works.* 2 h/week.

1892/93 – I.

Energetika *Energetics.* 5 h/week.

*A hőterjedés elmélete. (Publicum.) – *Theory of heat propagation* 1 h/week.

*Az előadásokkal kapcsolatos gyakorlatok – *Practical works.* 2 h/week.

1892/93 – II.

Az energia terjedési tüneményei – *Propagation of energy.* 5 h/week.

Az analitikus mechanika alaptanai (kezdőknek) – *Fundamentals of Analytical Mechanics (for beginners).* 2 h/week.

*Az előadással kapcsolatos gyakorlatok – *Practical works.* 2 h/week.

1893/94 – I.
Analitikus mechanika – *Analytical Mechanics.* 5 h/week.
*Az előadással kapcsolatos gyakorlatok – *Practical works.* 2 h/week.
1893/94 – II.
Erőfüggvények – *Force functions.* 5 h/week.
Fizikai mérőeszközök elmélete. (Publicum.) – *Theory of physical measure devices* 1 h/week.
*Az előadásokkal kapcsolatos gyakorlatok – *Practical works.* 2 h/week.
1894/95 – I.
Energetika – *Energetics.* 5 h/week.
*Az előadásokkal kapcsolatos gyakorlatok – *Practical works.* 2 h/week.
1894/95 – II.
Az energia terjedési tüneményei – *Energy propagation.* 5 h/week.
Mechanika kezdőknek –*Mechanics for beginners.* 2 h/week.
*Az előadásokkal kapcsolatos gyakorlatok – *Practical works.* 2 h/week.
1895/96 – I.
Mozgástan – *Cynematics.* 5 h/week.
Matematikai bevezetés az elméleti fizikába – *Mathematical introduction to Theoretical Physics.* 3 h/week.
Mozgástani gyakorlatok – *Practical works in Cynematics.* 2 h/week.
1895/96 – II.
Erőtan – *Force theory.* 5 h/week.
A mennyiségtani természettan művelése hazánkban. (Publicum) – *Mathematical Physics in our country.* 1 h/week.
*Erőtani gyakorlatok – *Force theory practice.* 2 h/week.
1896/97 – I.
Az energia alakváltozásai – *Energy form changes.* 5 h/week.
Matematikai bevezetés az elméleti fizikába – *Mathematical introduction to Theoretical Physics.* 3 h/week.
*Energetikai gyakorlatok – *Practical works in Energetics.* 2 h/week.
1896/97 – II.
Energia terjedési tüneményei – *Energy propagations.* 5 h/week.
Matematikai bevezetés az elméleti fizikába – *Mathematical introduction to Theoretical Physics.* 3 h/week.
*Elméleti fizikai gyakorlatok – *Practical works in Theoretical Physics.* 2 h/week.
1897/98 – I.
Általános mozgástan –*General Cynematics.* 5 h/week.
Matematikai bevezetés az elméleti fizikába – *Mathematical introduction to Theoretical Physics.* 3 h/week.
*Mozgástani gyakorlatok –*Practical works in Cynematics.* 2 h/week.
1897/98 – II.
Erőtan – *Force theory.* 5 h/week.
*Erőtani gyakorlatok – *Practical works in Force theory.* 2 h/week.
1898/99 – I.
Az energia átalakulásai – *Energy transformations.* 4 h/week.
Vektor-tan. (Matematikai bevezetés az elméleti fizikába) – *Vector theory (Mathematical introduction to Theoretical Physics).* 3 h/week.
*Energetikai gyakorlatok – *Practical works in Energetics.* 2 h/week.
1898/99 – II.
Az energia terjedése – *Energy propagation.* 4 h/week.
Az analitikus mechanika alaptanai – *Fundamentals of Analytical Mechanics.* 3 h/week.
Matematikai fizikai szeminárium – *Seminar in Mathematical Physics.* 2 h/week.
1899/1900 – I.

A mechanikai alapelvek némely alkalmazásai – *Some applications of principles of Mechanics*. 4 h/week.

Vektor-tan – *Vector theory*. 3 h/week.

*Mennyiségtani fizikai szeminárium – *Seminar in Mathematical Physics*. 2 h/week.

1899/1900 – II.

Erőtan –*Force theory*. 4 óra.

A mechanika alaptanai – *Fundamentals of Mechanics*. 3 h/week.

Matematikai fizikai szeminárium. (Tárgy: erőtan.) – *Seminar in mathematical physics. (Subject: Force theory)*. 2 h/week.

1900/1901 – I.

Az energia átalakulásai – *Energy transformations*. 4 h/week.

Vektor-tan –*Vector theory*. 3 h/week.

Mennyiségtani fizikai szeminárium. (Tárgya: energetika.) – *Seminar in Mathematical Physics. (Subject: Energetics)*. 2 h/week.

1900/1901 – II.

Az energia kérdése – *Energy problems*. 4 h/week.

Az elméleti mechanika alaptanai –*Fundamentals of Theoretical Mechanics*. 3 h/week.

Matematikai fizikai szeminárium –*Seminar in Mathematical Physics*. 2 h/week.

1901/1902 – I.

Különös mechanika –*Special Mechanics*. 4 h/week.

Vektor-tan – *Vector theory*. 3 h/week.

Matematikai fizikai szeminárium –*Seminar in Mathematical Physics*. 2 h/week.

1901/1902 – II.

Erőtan –*Force theory*. 4 h/week.

A mechanika alaptanai –*Fundamentals of Mechanics*. 3 h/week.

Matematikai fizikai szeminárium – *Seminar in Mathematical Physics*. 2 h/week.

1902/1903 – I.

Vektor-tan –*Vector theory*. 3 h/week.

Az energia átalakulásai – *Energy transformations*. 4 h/week.

Matematikai fizikai szeminárium – *Seminar in Mathematical Physics*. 2 h/week.

1902/1903 – II.

Az energia terjedése – *Energy propagation*. 4 h/week.

A mechanika alaptanai – *Fundamentals of mechanics*. 3 óra.

Matematikai fizikai szeminárium – *Seminar in Mathematical Physics*. 2 h/week.

1903/1904 – I.

Különös mechanika – *Special Mechanics*. 4 h/week.

Vektor-tan – *Vector theory*. 3 h/week.

Matematikai fizikai szeminárium – *Seminar in Mathematical Physics*. 2 h/week.

1903/1904 – II.

Erőtan – *Force theory*. 4 h/week.

Az analitikus mechanika alaptanai – *Fundamentals of Analytical Mechanics*. 3 h/week.

Matematikai fizikai szeminárium – *Seminar in Mathematical Physics*. 2 h/week.

1904/1905 – I.

Vektor-tan – *Vector theory*. 3 h/week.

Az energia átalakulásai – *Energy transformations*. 4 h/week.

Matematikai fizikai szeminárium – *Seminar in Mathematical Physics*. 2 h/week.

1904/1905 – II.

Az energia terjedése – *Energy propagation*. 4 h/week.

A mechanika alaptanai – *Fundamentals of Mechanics*. 3. h/week.

Matematikai fizikai szeminárium – *Seminar in Mathematical Physics*. 2 h/week.

1905/1906 – I.

Különös mechanika – *Special Mechanics*. 4 h/week.

Vektor-tan – *Vector theory*. 3 h/week.

Matematikai fizikai szeminárium – *Seminar in Mathematical Physics.* 2 h/week.
1905/1906 – II.
Erőtan – *Force theory.* 4 h/week.
Az analitikus mechanika alaptanai – *Fundamentals of Analytical Mechanics.* 3 h/week.
*Matematikai fizikai szeminárium – *Seminar in Mathematical Physics.* 2 h/week.
1906/1907 – I.
Vektor-tan – *Vector theory.* 3 h/week.
Az energia átalakulásai – *Energy transformations.* 4 h/week.
Matematikai fizikai szeminárium – *Seminar in Mathematical Physics.* 2 h/week.
1906/1907 – II.
Az energia terjedése – *Energy propagation.* 4 h/week.
A mechanika alaptanai – *Fundemantals of Mechanics.* 3 h/week.
Matematikai fizikai szeminárium – *Seminar in Mathematical Physics.* 2 h/week.
1907/1908 – I.
Analitikus mechanika – *Analytical Mechanics.* 4 h/week.
Vektor-tan – *Vector theory.* 3 h/week.
Matematikai fizikai szeminárium – *Seminar in Mathematical Physics.* 2 h/week.
1907/1908 – II.
Erőtan – *Force theory.* 4 h/week.
Mechanika alaptanai – *Fundamentals of Mechanics.* 3 h/week.
*Matematikai fizikai szeminárium – *Seminar in Mathematical Physics.* 2 h/week.
1908/1909 – I.
Vektor-tan – *Vector theory.* 3 h/week.
Az energia átalakulásai – *Energy transformations.* 4 h/week.
Matematikai fizikai szeminárium – *Seminar in Mathematical Physics.* 2 h/week.
1908/1909 – II.
Az energia terjedése – *Energy propagation.* 4 h/week.
Mechanika alaptanai – *Fundamentals of Mechanics.* 3 h/week.
Matematikai fizikai szeminárium – *Seminar in Mathematical Physics.* 2 h/week.
1909/1910 – I.
Analitikus mechanika – *Analytical Mechanics.* 4 h/week.
Vektor-tan – *Vector theory.* 3 h/week.
Matematikai fizikai szeminárium – *Seminar in Mathematical Physics.* 2 h/week.
1909/1910 – II.
Erőtan – *Force theory.* 4 h/week.
Mechanika alaptanai – *Fundamentals of Mechanics.* 3. h/week.
Matematikai fizikai szeminárium – *Seminar in Mathematical Physics.* 2 h/week.
1910/11 – I.
Vektor-tan – *Vector theory.* 3 h/week.
Az energia átalakulásai – *Energy transformations.* 4 h/week.
Matematikai fizikai szeminárium – *Seminar in Mathematical Physics.* 2 h/week.
1910/11 – II.
Az energia terjedése – *Energy propagation.* 4 h/week.
Mechanika alaptanai – *Fundamentals of Mechanics.* 3 h/week.
Matematikai fizikai szeminárium – *Seminar in Mathematical Physics.* 2 h/week.
1911/12 – I.
Analitikus mechanika – *Analytical Mechanics.* 4 h/week.
Vektor-tan – *Vector theory.* 3 h/week.
Matematikai fizikai szeminárium – *Seminar in Mathematical Physics.* 2 h/week.
1911/12. – II.
Erőtan – *Force theory.* 4 h/week.
Mechanika alaptanai – *Fundamentals of Mechanics.* 3 h/week.
Matematikai fizikai szeminárium – *Seminar in Mathematical Physics.* 2 h/week.

1912/13. – I.
Vektor-tan – *Vector theory.* 3 h/week.
Az energia átalakulásai – *Energy transformations.* 4 h/week.
*Matematikai fizikai szeminárium – *Seminar in Mathematical Physics.* 2 h/week.
1912/13. – II.
Az energia terjedése – *Energy propagation.* 4 h/week.
Mechanika alaptanai – *Fundamentals of Mechanics.* 3 h/week.
Matematikai fizikai szeminárium – *Seminar in Mathematical Physics.* 2 h/week.
1913/14. – I.
Analitikus mechanika –*Analytical Mechanics.* 4 h/week.
Vektor-tan – *Vector theory.* 3 h/week.
Matematikai fizikai szeminárium – *Seminar in Mathematical Physics.* 2 h/week.
1913/14. – II.
Erőtan – *Force theory.* 4 h/week.
A mechanika alaptanai – *Fundamentals of Mechanics.* 3 h/week.
Matematikai fizikai szeminárium – *Seminar in Mathematical Physics.* 2 h/week.
1914/15. – I.
Vektor-tan – *Vector theory.* 3 h/week.
Energia átalakulásai – *Energy transformations.* 3 h/week.
Matematikai fizikai szeminárium – *Seminar in Mathematical Physics.* 2 h/week.
1914/15. – II.
Mechanika – *Mechanics.* 3 h/week.
Optika – *Optics.* 3 h/week.
Matematikai fizikai szeminárium – *Seminar in Mathematical Physics.* 2 h/week.

---

* Seminars and courses exempt from charges.

The Development of the Cluj/Kolozsvár School of Mathematics (A hundred years ago, even more...)

József KOLUMBÁN*

Faculty of Mathematics and Computer Science
Babeş-Bolyai University
RO 400084 Cluj-Napoca
Str. Kogălniceanu 1

Abstract. In this paper we deal with the development of the School of Mathematics in Cluj (Kolozsvár) at the ending of the XIX$^{\text{th}}$ century and beginning of the XX$^{\text{th}}$ century.

AMS 2000 subject classifications: 01A70
Key words and phrases: Franz Joseph Univesity, higher education

1 Prologue

At the centurial commemoration of János Bolyai's birth in Cluj (Kolozsvár) the board of Hungarian Academy of Sciences announced the foundation of the Bolyai prize, intending to award it in five-year periods to a mathematician with outstanding achievements during the previous years. The person who first was honored with this prize was the French mathematician, Henri Poincaré. With respect to this event Jenő Gergely, who once had been the disciple of Frigyes Riesz, then professor at the Bolyai University, related the following story.

Many prominent representatives of the Hungarian scientific life waited in excitement the arrival to the Eastern Railway Station in Budapest of the Paris express. One of the greatest scientists of those times was coming to Budapest to receive the Bolyai prize. After the welcoming speech Poincaré asked: "Where is Feujé", he asked. The Hungarians looked at each other in confusion. Who is Feujé? Suddenly they realized that Poincaré was talking about Lipót Fejér, a professor at the Cluj (Kolozsvár) university, who, despite of his young age (he was only 25 years old at that time), was one of the most famous Hungarian mathematicians of those times. His results related to the trigonometric series had been published in Comptes Rendus (Paris) when he had been only twenty years old. Fejér's summation method has proved to be the starting point of the renaissance of Fourier's series. Fejér had started simultaneously his research enriching significantly the knowledge of the classical polynomials. He deduced a new and probably the simplest proof of Weierstrass' approximation theorem, and started his examinations on Chebyshev's polynomialss; these examinations had a great echo. Since then the main periodicals published regularly important

* email: kolumban@math.ubbcluj.ro

results achieved by Fejér. Thus we can easily understand Poincaré's wish to meet the young scientist during his short visit to Hungary. What was to be done? Thanks to an efficient mediation a train consisting only of an engine and one carriage started to rush within a few hours from Cluj to Budapest...

Whether it is true or not, the story is interesting, also because it reflects the condition of mathematics in Cluj at that time. Lipót Fejér was not the only mathematician, whose activity had a great influence on the scientific development in the 20th century, and who was working for a while at the university of Cluj. It is remarkable that it was the university of Cluj, not that one in Budapest among Hungarian universities that first became an important centre of mathematics. During the decades preceding to the world war the school of mathematics in Cluj became one of the best centres of mathematics in the world. Its representatives gained imperishable merits for themselves and for the Hungarian scientific life. It is not probable that a week day passes unless mentioning somewhere in the world the name of Gyula Farkas, Lipót Fejér, Alfréd Haar or Frigyes Riesz.

How did mathematics in Cluj reach so high at the beginning of the 20th century? What personal conditions, social and economic facts, strategies aiming the scientific life and the development of education contributed to this development? (The answer to these questions could be helpful when handling the present problems of the Hungarian minority in Transylvania.)

2 The first decades

In 1872, when the university of Cluj was established the situation of mathematics wasn't promising at all. Apart from the activity of the Bolyais the mathematics literature of that time was quite poor here. The main event worthy to mention is related to Sámuel Brassai, who translated into Hungarian the book of Euclid called Elements.

When the departments of mathematics had to be organised within the newly established university of Cluj, only Brassai could have come into consideration among the Transylvanians. At that time he was the director of the Transylvanian Museum Society and the guard of the Museum of Natural Sciences. His scientific reputation and prestige was so high in Cluj that the articles on the university repeatedly mentioned him as a nominee. He competed in 1870 for a teaching position at the university in Pest, and expected to be nominated as a professor of the Sanskrit language. As far as we know, the minister of education, József Eötvös asked Brassai if he was wanting to direct a department at the new university in Cluj. He was free to choose, but he indicated several domains, and this caused later problems to the new minister of education, Trefort securing the nominations, as "it was hard to agree on the department to nominate him at". Brassai proposed to choose between the philosophy, botany, pedagogy, history of culture, linguistics and mathematics. Finally he was nominated as a full professor at the Department of Elementary Mathematics. According to the book of György Boros entitled The Life of dr. Sámuel Brassai he himself was much surprised by this nomination..

Sámuel Brassai (1800–1897) was born in Colţeşti (Torockószentgyörgy). He graduated his studies in philosophy at the Unitarian College in Cluj when he was twenty-one years old. (He didn't study at any university.) He was an educator for a few years, and then in 1837 he became teacher of geography and history at the Unitarian College in Cluj, later he taught mathematics and natural sciences. He peregrinated to Germany in 1841, and returned to the college as director. After the 1848 Hungarian War of Independence he worked as a teacher in Pest, then again in Cluj.

The translation of Elements by Brassai still can be found in the mathematics library in Cluj. From his many coursebooks only the Algebraical Exercises published in 1883 in Budapest can be found in Cluj. One who is interested in the development of the mathematical terminology can learn a lot from these works. Unfortunately he didn't recognize the importance of the epoch-making discovery of János Bolyai, and his public statements impeded its recognition in Hungary. The other appointed full professor of mathematics was Lajos Martin at the Department of Higher Mathematics, who became the corresponding member of the Hungarian Academy of Sciences as an engineer in 1859.

Lajos Martin (1827–1897) was born in Budapest, he studied at the Technical University of Budapest and at the Genie Academy. He was an officer of the engineer corps until 1859, then he became the chief engineer of Buda in 1861, and was a high-school teacher between 1863–68. As a professor, due to the assignment of the minister of education he wrote coursebooks in mathematics, geometry and descriptive geometry. In 1868 he was the warden of the telegraph office of Pest, in 1869 the acting director of the telegraph office of Debrecen. From there he was nominated in 1871 as a director of the telegraph office of Cluj. Since 1872 until his death he was a full professor at the Higher Mathematics Department of the Cluj university. His scientific activity focused on the issues of aviation: he was the Hungarian pioneer of aviation. His main works: *A vízszintes szélkerék*, Budapest, 1874 (The Horizontal Wind Wheel), *Az erőmûtani csavarfelületek*, Budapest, 1874 (The Power Station Helicoidal Surfaces), A csillagászat újabbkori haladásáról, Cluj, 1877 (Recent Developments in Astronomy), *Variátio Számítás*, Cluj, 1879 (Calculus of Variations), *A madárrepülés általános elmélete*, Cluj, 1890 (A General Theory of Bird Flight). The first two works are bound in a single volume and the Calculus of Variations still can be found in the mathematics library in Cluj.

In the strict sense of the word neither Brassai, nor Martin were mathematicians. None of them performed mathematical research; mathematics wasn't in the centre of their interests. However both were interested in applied mathematical issues: the former confirmed this interest by his works in astronomy, the later by his works in aviation technology. The proper person for directing the third mathematics department, the Department of Mathematical Physics (*Mennyiségtani Fizika*) was found only in 1874, this person being the young Mór Réthy. He was one of the extremely talented young mathematicians, who acquired thorough knowledge and doctoral degree at famous German universities and who started to work at the university of Cluj. They gave a new stimulus to the local scientific life. **Mór Réthy** (1848–1925) was born in Nagykőrös, and studied at the universities of Vienna, Goettingen and Heidelberg. Taking his doctorate in Heidelberg in 1874, he became in that year extraordinary professor at the university of Cluj, after two years a full professor at the Department of Mathematical Physics. He was the head of the Department of Elementary Mathematics between 1884–1886. After that he moved to the Technical University in Budapest. He was one of the first Hungarian professors in theoretical physics. His results are well known abroad too, they relate to the shape of the incompressible jet. He performed significant research related to the principles of mechanics and the Ostwald rule in chemistry. His results relating to the infinitely equal areas examined by Farkas Bolyai are outstandingly valuable in the field of mathematics. It was Mór Réthy together with Gyula Kőnig who edited the first volume of the second edition of Farkas Bolyai's main work, the *Tentamen.*

In 1874 he presented a lecture in Cluj on Bolyai's geometry; this was the first lecture on this issue in Hungary. (We note here that the excellent Austrian mathematician, Johannes Frischauf already had course-like lectures in the academic year of 1871–72 at the university of Graz on the non-Euclidean geometry. These lectures comprised the first detailed overview of János Bolyai's brilliant work, the Appendix. The lectures were published in Leipzig in 1872 under the title Absolute Geometrie nach Johann Bolyai. This booklet was the only treaty for a long time, which constructed the absolute geometry using the elementary method, relying on the *Appendix.* However the book doesn't mention the construction problems that János Bolyai solved in the hyperbolic space.)

Réthy tried to ease the study of the Appendix by ensuring the renaming of several definitions (i.e. parallelism, paracycle, parasphere, hypercycle and hypersphere as interpreted by Bolyai), which were easier to understand than the original ones. Réthy built up individually the Bolyai trigonometry making use of the fact that the theorems of the Euclidean geometry apply for the infinitely small part of space in the absolute geometry, and relying on the fact that the trigonometry of the surfaces with a constant curvature is independent from the fifth Euclidean postulate. His other important achievement is that he was the first to appreciate and detail the constructions to be found in the Appendix. Réthy propagated the idea that the national mathematical research should rely on the results of the Bolyais. He said that "in our country, where besides the two Bolyais there weren't important mathematicians, all further scientific efforts should depart from the activity of these men."

The university had to face a lot of difficulties at the beginning. The teaching took place in several, mainly impractical buildings. The departments and the university library were in a deplorable situation. It is true that the Transylvanian Museum Society rented to the university for 50 years, at a price of five hundred forints a year all its collections, including its library of 30,000 volumes, but the special libraries serving education and scientific research were still mising. Compared to the needs the state could support the university at the beginning just with small amounts. The population of Cluj itself "was less enthusiastic, and endowed only four foundations on the behalf of the youth." The professors found themselves proper apartments only after long searching, and the students didn't find enough entertaiment in the small town, therefore "they found compensation in noisy amusements".

Considering these preliminaries no wonder that mathematics in Cluj had to make up the gaps for a couple of years after the establishment of the university. The difficulties were increased by the fact that in those years the whole Hungarian mathematics struggled with similar problems. The most competent persons knew well that the university could fulfil its real function if besides education high-level scientific research was also carried out within it. Despite this, it seems that during the first years there was a misunderstanding among the professors concerning the aims and forms of university activities. In the history and statistics of the Franz Joseph University of Cluj published in 1896 it is mentioned that the rector, Sándor Imre in 1878 raised objections against the lack of scientific research, and emphasized that the university wasn't a school, but a scientific institution, and one should not confuse these two. In the same time "considering that the main and most important task is to introduce and disseminate sciences, he urged firmly the reform of high-schools, so that young people should come properly trained to the university." When the new rector was introduced next year, the old one referred with pride to the competition ceremony held twice in the last year and to the signs of scientific diligence and independence in students, respectively to the work of the teaching staff, well-known abroad too, which contributed to the development of the national and universal scientific life. Whereas Sámuel Brassai, the new rector questioned on the importance that professors should not endeavour to new results in science alone, but should lead their disciples too in this direction. "It is not the first and main task of a university teacher to discover something new, and the demands for such research is useless, illegitimate and unreasonable. It is even more useless, illegitimate and unreasonable in relation to students." That is why he objected to prizes, as they being "the embodiments of the repeatedly refused demand, that the student should be directed towards new discoveries and scientific development."

Fortunately this idea wasn't prevalent within the university in the following years. The university paid more and more attention to the thorough training of the most talented students and to their introduction into research. The research competitions organised for students became more and more popular, and the curriculum reflected more and more the actual issues in science. The further training of the best students was supported through scholarships. In the meantime the management of the university devoted great care to assign professors, who acquired doctoral degrees at foreign universities, and had had scientific achievements there. These professors supported the development of talented students through several special courses that probed deeply into fashionable disciplines and comprised the often basic results of the lecturer. In Cluj the research focused mainly on number theory, the theory of differential equation, function theory, vector algebra, analysis, quaternion and elliptic function theory and the Bolyai geometry. The effectiveness and high scientific level is shown not only by the few published monographs, the countless lithographed notes, but also the increasing number of treaties published in domestic and foreign periodicals.

The state allows more and more funds for the university: during the first 15 years the costs increased from 160,000 forints to 279,000 forints. The press and the parliament appreciated the scientific activity of the professors and of the university. Despite all these efforts the number of students in mathematics decreased from 67 to 23 in the 1876-1886 period. The management of the university thought that its reason was that the eminent candidates waited in vain to be nominated as professors. Another reason could be extremely high prices in the town.

3 New impulses

At the beginning of the last decade of the 19th century important changes occurred in the university life. By that time the number of departments increased to 51, the members of the teaching staff to 68, while the number of the scientific employees comprising the teaching assistants and trainees to 119. During the first 25 years the number of students increased from 233 to 702. Similarly the number of mathematicians also increased.

In order to solve the building problems, the architect from Cluj, Károly Reményik started to build up the main building in 1893 according to the plans of Károly Meixner. Thus the rector, Lajos Martin, who had spoken about the airplane on the occasion of his inauguration, could welcome the youth in the new building in 1895. In the same year the municipality of Cluj gave the plot situated at the corner of Bel-Torda and Külső-Torda streets to the state in order to build there a modern library. By the turn of the century the university became the intellectual centre of Transylvania, the sanctuary of science, where — as count Imre Mikó expressed it — "the words of knowledge, enlightenment, and thus the words of freedom are advocated". All this is inseparable from the social, technical and economical development that characterised Hungary and Transylvania. The bourgeois society emerging in these years needed the development of education and science. More and more young people coming from bourgeois families were attracted to sciences. It happened quite often that a young person gave up his secure job and carrier for the sake of a scientific carrier.

We must mention the benefit of the Hungarian high-school reforms. The law on public education linked with the name of József Eötvös and the emancipation of Jews meant major changes in the field of mathematics and physics too. A modern network of schools was built up by the end of the nineties, where the education became demanding and effective. Excellent high-school teachers showed up, whose tireless work won many young people over to sciences.

The mathematics part of the modern Trefort curriculum elaborated for high-schools was worked out by Gyula Kőnig. He criticized the previous ineffective methods of mathematical education. He wrote: "We have to ask if one could expect better results on the basis of a methodology, which transforms the high-school mathematics curriculum into a series of abstract, uninteresting, thus incomprehensible «truths»; due to this methodology the students can hardly understand what serves all what they learn, and its aim is not to develop the mathematical thinking, but to teach as many «theorems» as possible."

Loránd Eötvös (1848–1919) and **Gyula Kőnig**(1849–1913) founded in 1891 the Society of Mathematics and Physics. Loránd Eötvös formulated the purpose of the Society as follows: "Let's learn from each other, thus we teach better." The teacher training colleges and training schools established on the basis of the concepts of Mór Kármán (the father of the world-famous physicist, Tódor Kármán) next to the universities in Budapest and Cluj (like the Demonstration School in Budapest) resulted in the increase of the teachers' training level. The specific Hungarian teacher-training model was completed with the establishment of the Eötvös College. In this college talented students could learn under the guidance of famous professors.

The "Hungarian wonder", the *Középiskolai Matematikai Lapok* (High School Mathematics Journal) launched by the young professor from Győr, Dániel Arany in 1893 played an important role in mathematics and physics training. Among such periodicals this is the journal lasting for the longest time in the world. (After three years the legendary teacher of the Lutheran College in Budapest, László Rátz took over the editing of the journal.) It turned out very soon how useful was when gifted young people were racking their brains on dozens of exercises for years and they wrote down their thoughts. Part of the students educated by the journal called today officially KöMaL became scientists, the others "only" very good professionals or professors. The best Hungarian mathematicians of the 20th century came among the students solving the exercises, like those who developed mathematics in Cluj before the First World War to the state-of-the-art.

Who were in those times the personalities whose work had an effect on the development of mathematics in Cluj? Which works can be still found in Cluj from those elaborated here besides those published in periodicals, despite the vicissitudes? Even a brief answer could be able to describe the local evolution of mathematics at the beginning of the new century.

We continue the enumeration with **Gyula Vályi** (1855-1913), the most famous disciple of Réthy. He was born in Marosvávsárhely (Târgu Mureş), and was the first important mathematician teaching at the university of Cluj, originated from Transylvania. He graduated his studies in 1877 in Cluj, then with the support of the university attended the lectures of Weierstrass, Kirchhoff, Kronecker, Borchardt and Kummer in Berlin for four semesters. He sustained his thesis entitled Contributions to the theory of second-order partial differential equations in 1880 in Cluj. He became an associate professor in 1881 at the university of Cluj. He was promoted full professor in 1884 of mathematical physics, then in 1886 of elementary mathematics. He published several treaties in the Journal of Mathematics and Natural Sciences. As Réthy mentions, "it is not the number and size of his treaties, but their quality, which is impressive." He was impeded in his work by his frail health; despite this "his spirit led him to create lasting things". According to the curriculum of the university it was Gyula Vályi who first presented course-like lectures on the Appendix in the second semester of the 1891–1892 academic year. He repeated this popular course almost invariably every four years. The lithographed lecture notes of 102 pages were published in 1904 in Cluj, with the title On János Bolyai's Appendix. He completed the original demonstrations by explanations. When comparing the absolute and the hyperbolic geometry he made use of Lobatschewsky's results in some places.

It is also due to Gyula Vályi's enthusiastic work that Cluj became the centre of the Bolyai cult, and many of his disciples had achievements in the further development of the Bolyai geometry. As a member of the board of the Transylvanian Museum Society, Gyula Vályi represented for many years the mathematicians.

Gyula Farkas and Lajos Schlesinger, who came to Cluj ten years later, were an outstandingly great asset to the university. Their activity had a predominant effect on the local mathematics. They used their professional prestige and aptitude to ensure the conditions of training and research in mathematics at the highest level and to secure the worldwide recognition of the results of Cluj school.

Gyula Farkas (1847–1930) was born in Sárosd. He studied at the university of Pest, where Ányos Jedlik had a great influence upon him. Later, thanks to count Géza Batthány, he peregrinated to France. Among his youthful achievements in mathematics we mention here only his examinations referring to Farkas Bolyai's trinom equation, which had referred to the algorithm for the approximation of the roots briefly presented in the Tentamen. Thus the Bolyai algorithm became well known, and many Hungarian and foreign mathematicians have been interested in its generalization, applicability, and the related convergence issues, even in our days. Gyula Farkas was an associate professor until 1887 in Pest, when he was nominated extraordinary professor at the university of Cluj. In 1888 he became a full professor at the university, he held this position until 1915. His former assistant, Rudolf Ortvay wrote on him: "his scientific activity, as well as his activity in public affairs related to the university was characterized by thorough criticism, the unyielding search for the truth that could not be diverted by issues of minor importance... And precisely because he didn't look for popularity, he acquired a great authority and had a fruitful influence on the direction of university affairs."

As a professor his interest focused mainly on problems related to theoretical physics, but he elaborated the background of the examined physical problems in such deepness that there are classical mathematical results as well among these elaborations. Dating from the nineties he was particularly preoccupied by the Fourier principle of mechanics. His studies specify the necessary condition of balance in case of conditions given by inequalities. For this he demonstrates his theorem on linear inequality, which is one of the most famous Hungarian mathematical achievements, known as the Farkas-lemma. On the basis of these works it is obvious today that Gyula Farkas was one of the modern creators of the optimization theory.

He published his university courses in carefully elaborated lithographed publications. His lecture notes can be still found in the mathematics library in Cluj, i.e.: *Analytikus mechanika*, 1907–08 (Analytical Mechanics), *Analitikus mechanika*, 1913–14 (Analytical Mechanics), *Erőtan*, 1913–1914 (Dynamics), A mechanika alaptanai, 1913-14 (The Principles of Mechanics). The *Vector-tan és az egyszerû inaequatiok tana* (Theory of Vectors and the Theory of Simple Inequalities) published in Cluj is also accessible. This one is a well-written book on vector analysis comprising the main results of his research.

Lipót Klug (1854–1944) was born in Gyöngyös. He took a diploma in mathematics at Budapest. He was a professor in Pozsony (Bratislava), then in Pest. Between 1897–1917 he was a professor at the department of descriptive geometry at the university of Cluj. Meanwhile he published five popular coursebooks in addition to numerous treaties: *A projektív geometria elemei*, Budapest, 1892 (The Elements of the Projective Geometry), *Projektív geometria*, Budapest, 1903 (Projective Geometry), *Az általános és négy különös Pascal-hatszög configuratiója*, Budapest, 1898 (The Configuration of the General and Four Special Pascal hexagons), *Ábrázoló geometria*, Budapest, 1900 (Descriptive Geometry) and *A harmadrendû térgörbék synthetikai tárgyalása* (Synthetic Treaty on the Curvature of the Third Order). The first four of these still can be found in Cluj. Unfortunately only one lecture notes remained: *Az egyszerû görbe felületek ábrázolása*, 1909–10 (The Description of Simple Curvatures).

Lajos Schlesinger (1864–1933) was the best known professor abroad at the turn of the century. He was born in Nagyszombat (Trnava), he attended high school in Hungary, then studied in Heidelberg and Berlin. He taught in Cluj for a semester in 1890 as an associate professor of the Berlin university. He was the disciple and the son-in-law of the famous professor from Berlin, Lazarus Fuchs, that is why people said about him that "the talent is inherited mainly by the son-in-law".

In 1897, when he became a full professor at the university of Cluj, he already was a well known scientist. The Teubner Publishing House in Leipzig had published the first two volumes of his famous book (*Handbuch der Theorie der Linearen Differentialgleichungen*) that he had written as the associate professor of the Berlin university. He published the third volume in 1898 as a professor in Cluj. The society of mathematicians knew him by that time as one of the prominent authorities in the theory of differential equations built on the complex analysis. Many parts of this significant theory were enriched by Schlesinger.

After arriving to Cluj he got involved in education with a great enthusiasm besides the continuation of his research. Fifteen of his lecture notes still can be found in the mathematics library in Cluj. It might not be useless to enumerate their titles: *Elliptikus függvények elmélete és alkalmazásai*, 1898–99 (The theory of elliptic functions and its applications), *Égi testek mechanikája*, 1898–99 (The mechanics of celestial bodies), *A differentiál-számítás*, 1900 (The differential calculus), *Riemann-féle felületek*, 1900 (The Riemann surfaces), *Elliptikus függvények*, 1901 (Elliptic functions), *Bevezetés a variatio számításba*, 1902 (Introduction to the variation calculus), *A tér abszolute igaz tudománya* (The absolute true science of space, jubilary lectures on the 100th anniversary of János Bolyai), *Differenciálszámítás*, 190? (Differential calculus), *Az abszolút sík eltolásaiból alkotott discontinuus csoportokról*, 1905 (On the discontinuous groups formed from the translation of the absolute plane), *Fucks-féle függvények*, 1906-07 (Fucks functions), *Égitestek mechanikája*, 190? (The mechanics of celestial bodies), *Görbevonalak és felületek elmélete*, 1907-08 (The theory of curve lines and surfaces), The theory of curve lines and surfaces *Válogatott fejezetek az infinitesimális geometriából*, 1908 (The theory of curve lines and surfaces), *Égi testek forgásáról*, 1908–09 (On the rotation of celestial bodies), *Differenciálegyenletek elmélete*, 1909-10 (The theory of differential equations). These lectures are fascinating due to their clear, distinct style and the accurate treatment of the referred latest results. There is no doubt that the above-mentioned lectures belong to the best results of mathematical education in Cluj.

The book entitled *Vorlesungen über lineare Differentialgleichungen* and published in 1908 by the Teubner Publishing House can be considered an important result of Lajos Schlesinger's scientific activity carried out in Cluj. This is not the revision of the theory elaborated in the above-mentioned Handbuch, but a treatment of linear differential equation systems using totally new methods. This is the first monograph treating the linear differential equation systems with variable coefficient using the product integral as interpreted by the Italian mathematician, V. Volterra.

It can be considered a recognition of his authority that the Teubner Publishing House published in 1909 his book entitled Bericht über die Entwicklung der Theorie der linearen Differentialgleichungen seit 1865, edited by the German Mathematics Society. He wrote his famous book, Automorphe Funktionen that still can be found in the mathematics library in Cluj as a professor at the university of Giessen.

During his stay in Cluj Schlesinger contributed significantly to the advancement of the local mathematics. Together with Gyula Farkas and Gyula Vályi he had a basic role in the establishment of an excellent mathematics library within the university.

4 The Bolyai centenary

In 1894 the Congres international de bibliographie des sciences mathematiques presided by Henri Poincaré was preparing a considerable bibliographical volum, which would have included according to the original plans a chapter entitled Geometrie de Lobatschewsky. Thanks to the mediation of a group of Hungarian scientists the title of the chapter was changed to Geometrie de Bolyai et Lobatschewsky. From this time on the scientific literature mentions these two names as equivalents in relation to the non-Euclidean geometry. Why so late? The recognition of János Bolyai's merits was delayed by several reasons. First Gauss, though he recognized the geniality of Bolyai ("Ich halte diesen jungen Geometer v. Bolyai für ein Genie erster Grösse", he wrote in one of his letters), unfortunately had a disadvantageous effect on the posterity's opinion on the Appendix. Gauss stated that it had been him who raised first the idea of the non-Euclidean geometry during his conversations with Farkas Bolyai, and that the father had mediated his idea to János. A similar opinion was expressed in relation to Lobatschewsky, in this case Bartels (the former disciple of Gauss, who later became a professor in Kazan) was considered the mediator of the new ideas. The truth is that János Bolyai discovered alone and independently his geometrical system as his heritage revealed. However the delay of the recognition can be attributed first of all to the neglect of the Hungarian mathematicians of those times. (It is typical that Sámuel Brassai criticises Bolyai's work even in 1886.) It is well known that the pioneers in the research on Bolyai's work were foreigners. Hoüel, a professor from Bordeaux translated the Appendix into French in 1867 and published it with a biography of Bolyai written by Ferenc Schmidt, an architect from Temesvár (Timisoara). The Italian translation edited by Battaglini was published in the same year, and the review in German, as we already mentioned, was published in 1872. In 1891 the English version was published thanks to the contribution of Halsted, a professor from Texas. So the representatives of the Hungarian scientific life recognized their debt. Under the effect of the foreign initiatives after a long wrangling the Appendix was finally published in Hungarian too, in 1897. These were the preliminaries preceding János Bolyai centenary.

The centenary celebrations were prepared in Cluj in accordance with the importance of the event. The Faculty of Mathematics and Natural Sciences of the university decided on its session held on December 29th, 1899, that on his hundredth anniversary the house where János Bolyai had been born would be supplied with a memorial plaque, a ceremonial commemoration would be held, and a Festschrift would be issued, which would outline the influence of the Bolyai geometry on the development of mathematics in the 19th century. Gyula Farkas and Lajos Schlesinger played the main role in the organisation of the centenary celebrations.

Lajos Schlesinger identified the birth house of Bolyai. This house is one of the famous buildings of Cluj today, thus it might be of interest if we touch upon the condition of its identification. A letter of Farkas Bolyai to Gauss written on September 11th, 1802, in Domáld gave the first point of reference in the research, where he specified the address as follows:

"Meine Adresse: Mr. Wolfg. Bolyai
Bodor Pál úrnal
a belső közép utczában"
("My address: Mr. Wolfg. Bolyai, Mr. Pál Bodor's house, in the Central street").

The following letters contain the same address until September 16th, 1804, when Farkas specifies his address in Marosvásárhely (Târgu Mureº). It seems possible that the Bolyai family during its stay in Cluj, namely between autumn 1802 and spring 1803 lived in Bodor's house in the Belközép street, and János was born in that house. However, according to the evidence given by many inhabitants in Cluj (i.e. László Bodor, judge of the Court of Appeal, the great-son of Pál Bodor) and by the letters written by Farkas Bolyai to Pál Bodor, and preserved in the archives prove that the parents of Mrs. Bolyai (born Zsuzsanna Benkő) had their own house. Pál Bodor sold this house by auction in 1816 to Jakab Szenkovits on behalf of family the widow of József Benkő. Schlesinger came to the conclusion that the Bolyai family had lived in the house situated also in the Belközép street, belonging then to József Benkő, later to Szenkovits, and that János Bolyai had been born in that house. Most probably the fact that Farkas Bolyai directed his letters when staying in Cluj, as well as when staying in Domáld to his friend, Pál Bodor can be explained by that he found useless to trouble Gauss with a change of address for such a short time, as his house in Cluj was so close to Pál Bodor's house, that he could gather his letters without any loss of time. On the basis of the letters of Bolyai found in Bodor's archives Schlesinger states firmly: "Thus it is ascertained that the house in question is the house located on the corner of the Tivoli and Belközép streets in the main square's direction, which now belongs to the merchants' society; as it has only one entrance from the Tivoli street, this house doesn't have a number on the Deák Ferencz street, but its address is Tivoli street no. 1. Thus it is proved that the house on Tivoli street no. 1 is the house where János Bolyai was born." [N.B. Belközép street got the name of Deák Ferenc in 1899.]

The Festschrift with beautiful leather binding written in Latinis entitled: *Libellus post saeculum quan Joannes Bolyai de Bolya anno MDCCCII a.d/ XVIII Kalendas Januarias Claudiopoli natus est, ad celebrandum memoriam eius immortalem ex consilio Ordinis Mathematicorum et Naturae Scrutatorum Regiae Litterarum Universitatis Hungariae Francisco-Josephinae Claudiopolitanae editus*, Claudiopoli, MCMII.

Besides the Latin translations of three famous letters written by János Bolyai to his father in 1823 from Timisoara (Temesvár) the book comprises three studies; the first two were published in 1903 in the Acta Universitatis (in Hungarian). These are the following:

– Lajos Schlesinger (Cluj): On the application possibilities of the absolute geometry on the complex variable function theory

– Paul Stäckel (Kiel): On the mechanics of multiple dimensional manifolds

– Robert Bonola (Pavia): The list of studies on absolute geometry published between 1839 and 1902.

The university board decided in November 1902 to celebrate János Bolyai's hundredth birth-anniversary on January 15th, 1903, considering that the guests coming from a great distance would have difficulties in travelling in the middle of December. They formed a committee to organise the celebrations, its four members were Dénes Szabó (president), Gyula Farkas, István Apáthy and Lajos Schlesinger. They invited Lajos Schlesinger to give a memorial speech. The Hungarian Academy of Sciences lead by Loránd Eötvös appointed a delegation on November 22nd, 1902, its members being Kálmán Szily, Mór Réthy, József Kürschák and Béla Tötösy.

On the assigned day, at ten o'clock in the morning the assembly hall of the new university building was already full with the delegates of the scientific institutions and societies, with the eminent officials and intellectuals of Cluj, with "a charming female audience" and youth. The Rector who was a historian opened the ceremonial meeting according to the style of the époque: "On the 15th of December 1802 a tiny star passed over our town's horizon; during its short career this star glittered as first-class brightness on the sky of our culture and of the universal science, in order to give light where humans have never looked at before. János Bolyai disclosed those secrets that had lead many brilliant talents to maze for 2000 years."

The presiding rector invited all the speakers of the delegations, who gave their welcoming speech as follows: Loránd Eötvös, the president of the Hungarian Academy of Sciences; dr. Izidor Fröhlich, representing the Hungarian Royal University; dr. Manó Beke representing the Faculty of Arts of the University of Budapest; dr. Gusztáv Rados from József Royal Hungarian Technical University from Budapest, Emánuel Budisavlievic lieutenant-colonel representing the Imperial and Royal Technical and Military Academy of Vienna; dr. József

Kürschák representing the Mathematics and Physics Institute; Lajos Csíki representing the Reformed College of Marosvásárhely (Târgu Mureº); dr. János Szamosi representing the Transylvanian Literary Society, and János Bedőházi, the writer of the two Bolyais' biography. The speeches made a deep impression on the audience, which expressed its appreciation by enthusiastic ovation. Besides the speakers many institutes and associations sent their delegates. With the occasion of the ceremony the Mathematics Institute in Goettingen sent its greetings via telegram, the Association of German Mathematicians and the Mining and Forestry Academy from Selmecbánya via letters.

As mentioned in the Prologue, the establishment of the Bolyai-prize was announced during the centenary celebrations. The initiators were Gyula Kőnig, Gusztáv Rados, József Kürschák, Gyula Farkas and Lajos Schlesinger. Kálmán Szily, the secretary-general of the Hungarian Academy of Sciences submitted the following report with reference to the prize: "The Hungarian Academy of Sciences contributes to the celebration of the hundredth anniversary of János Bolyai by deciding to establish a prize in memory of the immortal scientist, respectively of his father of profound thinking and master to his son in sciences. The Bolyai-prize consisting of 10,000 crowns and a medal will be awarded to the author of the most outstanding treaty in mathematics published within the last 5 years wherever and in whatever language, taking into account the previous scientific activity as well of the person in question. The prize will be awarded for the first time in 1905, after that in every five year during the meetings held in December. One side of the medal will show the building of Hungarian Academy of Sciences from Budapest, the other will contain a Hungarian inscription."

Unfortunately the prize could have been awarded only twice. During the First World War the international scientific relationships broke off, the money lost its value, thus such a proper initiative didn't have any continuation. As shown by his lecture notes, the Bolyai geometry was in the centre of the preoccupations of Lajos Schlesinger. Therefore it wasn't accidental that he was invited to give the memorial speech during the celebrations. He thoroughly prepared it. He enumerated in his meaty speech important arguments related to the priority debate on Bolyai and Lobatschewsky, thus he contributed significantly to the recognition of Bolyai's genius.

Following the memorial speech a few copies of the Festschrift were distributed, then at the rector's request the celebrating audience visited the house where János Bolyai had been born. They all wanted to be present when the representatives of the Faculty of Mathematics and Natural Sciences unveiled the memorial plaque placed on the facade on the Deák Ferenc street. During the unveiling Gyula Farkas (dean at that time) gave a speech. He said for example that "We call János Bolyai on this plaque the Hungarian Euclid, as he was a master in geometry like Euclid. We mention his father, Farkas Bolyai too, because he deserves this as the deep thinking author of the *Tentamen*, besides he was father in developing the talent in mathematics of János."

The text written on the plaque is as follows: "*Az 1802. év 12. havának 15. napján, itt született Bolyai János, a Magyar Euklides, bolyai Bolyai Farkasnak, a Tentamen mély gondolkodású szerzőjének fia, minek az emlékezetére száz év múltán a Ferencz József Tudományegyetem matematikai és természettudományi kara állítá e követ.*" ("In this place János Bolyai was born on the 15th of December, 1802, the Hungarian Euclid, son of Farkas Bolyai, the author of Tentamen with a profound thinking, in memory of whom a hundred years after the Faculty of Mathematics and Natural Sciences placed this stone.")

Simultaneously to the establishment of the Bolyai prize the Academy announced a call for writing a monograph on the Bolyai-Lobatschewsky geometry. We don't know about the results of the announcement. In 1902 Lajos Schlesinger held lectures at the university of Cluj under the title The absolute true science of the space, and according to Barna Szénássy this was intended to be the basis of the monograph prepared for the announcement. The partly unreadable manuscript still can be found in the mathematics library in Cluj. Its publication would be extremely important for the history of science.

It is certain that the centenary celebrations organised in Cluj meant the definitive victory of János Bolyai not only abroad, but in his own country too.

5 The peak

Besides his up-to-day valid scientific activity, Gyula Farkas had a predominant role in the development of the Mathematics Seminar in Cluj. Started out almost from nothing, the institute became after a quarter of century one of the best scientific workshops of the Monarchy. Due to his commitment to sciences and outstanding virtues Farkas had a great authority among his colleagues and students, which he made use of in the university administration. He was dean for seven times and once even the rector of the university. His word was always decisive. He was exigent, he expected much from himself and others. He appreciated the human values, and helped a lot those deserving it. He made use of his influence to improve the material and personal conditions of the university activity. It was due exclusively to his intercession that such brilliant young mathematicians came to the university like Lipót Fejér (1905), Frigyes Riesz (1911) and Alfréd Haar (1912), who together with the elder professors raised the quality of the mathematics training to such a high level in the first decades of the 20th century, that this level, considering the international reputation of the professors, would be hard to surpass.

The following letter written by Gyula Farkas to Lipót Fejér demonstrates how deeply Gyula Farkas was concerned about the cause of the domestic mathematics and physics. (These letters were published by András Prékopa in his treaty *Farkas Gyula élete és munkásságának jelentősége az optimalizálás elméletében* - The life of Gyula Farkas and the significance of his work for the optimization theory; see Sándor Komlósi and Tamás Szántai eds.: *Új utak a magyar* *operációkutatásban. In memoriam Farkas Gyula* - New methods in the Hungarian operations research. In memoriam Gyula Farkas, Dialóg Campus Publishing House, Budapest-Pécs, 1999.)

After he managed to appoint Lipót Fejér to the university of Cluj, he wrote the following letter on July 9th, 1905:

Honourable Dr. Lipót Fejér,

Tégla street, House Kolozsi, Gödöllő

Dear Doctor!

I am happy to find out from your kind letter that you are spending your time in the best place, in family circle, and the Alma Mater could not directly profit of this time of yours. Though I know that even so our common home gets its part too, indirectly: either you are spending it by resting, either by working; in the first case you are gaining energy, in the second you gain and transmit knowledge... I give you my best regards; please transmit them to your relatives too. I think constantly with love to our past and future collaboration.
Yours sincerely, Gyula Farkas

In his letter dated October 1st, 1911 Gyula Farkas is yet concerned with bringing Frigyes Riesz to Cluj:

My Dear Learned Colleague!

I kindly ask you to communicate me the address of Frigyes Riesz, so that following the decision concerning the substitution I can directly apply to him.
Your sincere friend, Gyula Farkas

He wrote on September 3rd, 1911 on the back of a card:

Honourable Dr. Lipót Fejér

University full professor

Budapest, Vörösmarty street 19, ground floor, 1st apartment

The committee has approved my proposal yesterday, and recommends to invite Frigyes Riesz to substitute the ordinary department and to manage the seminar. The faculty session will be held the day after tomorrow.

The next letter is dated October 20th, 1911.

Honourable dr. Lipót Fejér

University full professor

Budapest, Vörösmarty street 19, ground floor, 1st apartment

Most Respected Dear Colleague!

Today our new colleague, Frigyes Riesz has arrived; the dean, Vályi and myself met him at noon in the dean's office, and entrusted him to lead the mathematics seminar too. He instantly wrote on our black board the announcement *of his lectures and seminars, he will start his lectures on the 23rd. All this reassured us deeply. I transmitted to the dean's office your notification on Haar in the very momnet I received it.*

With kind regards,
Your true friend, Gyula Farkas

As Lajos Schlesinger left the university in summer 1911, Gyula Farkas got interested in Alfréd Haar. He wrote the following letter on November 4th, 1911:

Honourable dr. Lipót Fejér
University full professor
Budapest, Nádor street 51, 5th floor

Dear and Respected colleague!

Our colleague, Frigyes Riesz has already started his work with a great enthusiasm in the mathematics seminar, too... I received a letter from Haar yesterday. It seems that the invitation addressed to him denominates me as the faculty referent of the competition. Therefore in my answer I decided to describe the essentials. Otherwise Haar's letter confesses such deep patriotism that I am no more afraid to loose him in favour of some universities. I wish you a happy life in your new apartment,

Truly yours,
Gyula Farkas

During their stay in Cluj one of the main tasks of Lipót Fejér, later of Frigyes Riesz was the managing of the Seminar in Mathematics. The seminar directed by Gyula Farkas, Lajos Schlesinger and Gyula Vályi was established in 1901. Its real manager was Schlesinger until 1905.

Who were the above-mentioned representatives of the new wave? **Lipót Fejér** (1880–1959) was born in Pécs. During his high-school studies he solved successfully the exercises published in the Mathematics Journal. In 1897 he was placed second at the competition organised by the Society of Mathematics and Physics. In the same year he was admitted to the mechanical engineering section of the Technical University of Budapest, but after a semester he moved to the so-called general section, as a student in mathematics and physics. There he attended mainly the lectures of Gyula Kőnig, József Kürschák and Gusztáv Rados, then he moved to the University of Budapest. During the academic year 1899–1900 he studied at the University of Berlin, where he attended the courses of L. Fuchs, G. Frobenius and H. A. Schwarz. Once returned home he spent the 1900-1901 academic year at the University of Budapest. He published in Comptes rendus his theorem on Fourier's series that made his name all at once worldwide known. He wrote his dissertation in mathematics on Fourier's series, in physics on the phenomena of diffraction. From September 1st, 1901 he was an associate professor at the University of Budapest, in spring 1902 he acquired the doctor's degree there. He spent the first semester of 1902–1903 in Goettingen, where he attended mainly the lectures of D. Hilbert and H. Minkowski. He spent the second half of the same year in Paris, attending the lectures of E. Picard and J. Hadamard. He had nine publications until 1905, three among these in the Comptes rendus, one in the Matematische Annalen. After such preliminaries, as Gyula Farkas expressed it, no one would have liked to "loose him in favour of some foreign universities". The university of Cluj ensured him the possibility of advancement. He was employed as an associate professor at the Department of Mathematical Physics (its head being Gyula Farkas). Besides the proofs of his brilliant talent in mathematics a reason for this could have been that he had learnt physics earlier, and he was interested in theoretical mechanics. Following his habilitation on June 23rd, 1905 at the Faculty of Mathematics and Natural Sciences of the Ferencz József University in Cluj with his thesis called Stability and instability examinations in the mechanics of mass point system his star was raising rapidly. In the first half of the 1905-1906 academic year he was the assistant lecturer of Lajos Schlesinger, in the second semester the associate professor of analysis and analytical mechanics, in September 1906 senior lecturer, and in 1911 extraordinary professor. He finished his activity at the university of Cluj in summer that year, as from September he became full professor at the University of Budapest. In 1908 he was elected a correspondent member of the Hungarian Academy of Sciences. During his stay in Cluj he published around 30 scientific treaties, mainly in famous journals of mathematics. His treaties relate to Fourier's series, theoretical mechanics and analytical functions.

Frigyes Riesz (1880–1956), who directed the Department of Higher Mathematics between 1912 and 1919, was born in Győr, and he was educated with an extreme care. He studied at the Grammar School of the Benedictine Order in Győr, then started his university studies at the Technical University of Zurich. However his vocation for science prevailed over the attraction to make a career as an engineer, thus he continued his studies from 1899 at the University of Budapest, and then spent a year in Goettingen. In Budapest the lectures of Gyula Kőnig and József Kürschák, in Goettingen the lectures of Hilbert and Minkowski had the greatest influence on him. In 1902 he was conferred a doctoral degree in Budapest. He published his first scientific discovery, the so-called Riesz-Fischer theorem when he was 27 years old. Later the identity of the two composition of the quantum mechanics, the wave mechanics of Schrödinger and the matrix mechanics of Heisenberg has been proved using this theorem. It is also linked to Riesz the widely applied discovery of the space of the integrable function on the pth exponent, the representation of the linear functionals defined on the set of the continuous functions in the form of the Stieltjes integral, respectively the definition of the dual of the Hilbert spaces. In 1908 he defined in his presentation held at the international mathematics conference in Rome the term of the topological space. Due to all these youthful discoveries he became famous before coming to Cluj.

From his lecture notes only three can be found in Cluj today: *Függvénytan*, 1911–12 (Function theory), *Fourier-féle sorok*, 1913–14 (Fourier's series) and *Függvénytan*, 1914–15 (Function theory). The book of Riesz published in Paris and entitled *Les systemes d'équations linéaires a une infinité d'inconnues* is also available; it had an important role in the development of the functional analysis. During his work carried out in Cluj he published several important treaties.

Frigyes Riesz is not only one of the excellent Hungarian mathematicians, but he is considered worldwide as one who had the greatest effect on the grounding and development of modern branches of mathematical analysis. The terms and methods he had introduced, the achievements related to him belong today to the classical material of the real function theory, the functional analysis and the general topology. These results were partly included in his famous book written jointly with Béla Szőkefalvy-Nagy entitled Leçons d'analyse fonctionnelle, which was translated into several languages, and is known worldwide.

The last young mathematician who came to Cluj from the present territory of Hungary was **Alfréd Haar** (1885–1933). He was born in Budapest; he graduated at the Lutheran High School, where László Rátz, the famous editor of the Középiskolai Matematikai Lapok (High School Mathematics Journal) was his mathematics teacher. During his high school studies Haar took part diligently in the editing of the journal. In autumn 1903 Alfréd Haar won the first prize at the Eötvös Loránd Mathematics Competition organised every year for those who had acquired high school degree the previous year. He studied in Budapest and Goettingen. He attended the lectures and seminars of Beke, Eötvös, Frölich, Kürschák, Rados, Scholtz in Budapest, respectively of Carathéodory, Hilbert, Klein, Minkowski, Prandtl, Runge, Schwarzschild, Voigt and Zermelo in Goettingen. He took his doctor's degree under the guidance of Hilbert. After that he was an associate professor at the Technical University in Zurich. In 1912 he was nominated to one of the physics departments at the university of Cluj, first as extraordinary professor, then in 1917 as full professor. His work had an acknowledged effect on the modern development of mathematics. His talent was linked with the conscientiousness of the real scientist. His research covered a wide area. His results related to the systems of orthogonal functions, to the variation calculus, to the singular integrals, to the theory of sets, the function approximation, the linear equations and the topological groups are famous even today. The terms Haar's basis, the function systems having Haar features, but especially the term Haar's integral are definitely part of the every-day tools of mathematicians. The present effect of Haar's results are well proved by the lectures kept at the centenary scientific meeting in 1985, which are collected in the volume entitled Alfred Haar Memorial Conference.

Alfréd Haar didn't publish any book, but he compiled several well-written notes, which contain many original details, and could be used today as university coursebooks. Unfortunately only four can be found in Cluj: Differential-Gleichungen (Goettingen, 1911), Algebra (Algebra, 1912-13), Determinánsok és quadratikus formák, 1912-13 (Determinants and quadratic formes), Számelmélet, 1915-16 (Number theory).

6 Cluj, a beautiful town

Finally let us speak about Cluj how it looked a hundred years ago. László Passuth in his autobiographical novel entitled Kutatóárok (Research ditch) relates a few memories on the atmosphere of the town. "I remember the town as bordered on one side by the Monostor hill, on the other side by the green area of Hóstát. Next to the infinite main street of Hóstát a few thousands of peasants lived as corpus separatum. Quite a lot of railway workers formed settlements, there were a few smaller factories... There were relatively many banks, religious institutions, but above all schools... The flooding of students coming from outside the town lasted from September to June, and ensured livelihood for many meal-houses, tailors, hosts. The wisdom was shared with law students on seminars, and the «nightlife» of the town was quite colourful. . . There were just a few industrial workers, and I don't remember feeling any social tension in the form of wage-demands, strikes or unemployment. The social ideas penetrated mainly via theoretical sources. . . certain university teachers became «radicals», . . . the influence of the Workers' Savings Bank was strong (Béla Kun worked there too), the coexistence of different religions ensured a certain patience that could not stand any inflexibility. As perhaps a town with such a diversified map of religions nowhere existed in Hungary at that time, where the colleges partly remained denominational colleges. The majority was Catholic, but a great part of the native inhabitants was Calvinist. The Lutherans were mainly Saxons. . . The only episcopate of the Unitarians was here... Two «Greek» churches served the Romanian congregations: a Greek Catholic and an Orthodox church. Before the war a patrician Jewish community lived in the town, leading a neolog lifestyle. The biedermeier town with its small intellectual volcanoes was in fact a provincial centre of gossiping, though the habitants didn't know each other as much as in other towns with fifty-sixty thousand of inhabitants. The reason for this was the constant moving of people, first of all among the middle class: a great number of clerks settled down and moved away, according to the laws of a clerk's existence."

The centre got its present aspect in those times, the masterpiece of János Fadrusz, the statue of king Matthias Corvinus was placed there in 1902.

The café of the New York Hotel was the centre of social and intellectual life. There were two separate long tables in the hall, where university teachers and aristocrats got on well together. They met each other very often; they were drinking and moving together to the artists' tables. The actors of the National Theatre, writers, journalist had their permanent, traditional tables. "There are in Cluj a few obstinate, incorrigible idealists. They have subtle taste and are inclined to arts. You can find among them painters, sculptors, architects, writers"

— the journal called *Ellenzék* (Opposition) wrote this on January 3rd, 1905. The specific, historic atmosphere of the town, in the meantime its intellectual avant-gardism could be felt on these meetings. According to the oral traditions Lipót Fejér was a regular customer in the New York Café. That's how he met, during a tarot game the mother of Passuth, who was of the same age. According to Passuth's relating the young professor asked the woman after the game why she wouldn't join to the university, and he proposed to prepare her in mathematics... Fejér's range of interests went far beyond mathematics. He loved music passionately, and he himself played well the piano. Musicians and writers appreciated his opinion as well as aestheticians and law philosophers. He met several times Endre Ady too in the café. The portrait of Ady dedicated with cordial words illustrates their friendship.

The Farkas street with its beautiful chestnuts was the common sanctuary of science and muses. At one end there was the main building of the university, next to it the old theatre built of stone, opposite to them there was the Piarist High School, at the other end of the street there was the Reformed High School and the Reformed Church without a tower. These buildings and their institution defined the atmosphere of this famous district. Our great mathematicians entered their workplace, the main building of the university from this street.

7 Epilogue

In 1919 the university of Cluj entered under Romanian authority, therefore most of the professors – like Alfréd Haar and Frigyes Riesz – moved to Szeged, where they laid the foundations of a new university. In the meantime the Hungarian "mathematicians producing machine" continued its operation. A short time after the war great scientists showed up, like Tódor Kármán, György Pólya, Gábor Szegő, John von Neumann and others. However they didn't start towards the eastern regions of the Carpathian Basin anymore, alike their colleagues a few years before, but towards... But this one is another story to tell.

Gyula Farkas as a Bolyai Researcher

Róbert OLÁH-GÁL* **

Facultatea de Matematică şi Informatică
Universitatea Babeş-Bolyai Cluj-Napoca
Colegiul Miercurea-Ciuc, Romania

Abstract. Gyula Farkas came to Transylvania and to the university of Cluj from Budapest (University of Pest) and it was during his deanship that the centenaty of János Bolyai's birth was celebrated. But it was also him to represent the Academy at the exhumation and second interment of the two Bolyais in Marosvás arhely (Târgu Mureş). These events naturally cast light upon the work of the two Bolyais. So Farkas knew that as a scientist he is due to search for and preserve every document linked to the work of the two Bolyais. It is only natural that he performed this work of research in Kolozsvár (Cluj) where he spent the most of his life as a university lecturer and professor. One colleague of his in Kolozsvár, the similarly well learned Lajos Schlesinger, had just finished his research on the birthplace of János Bolyai at the same time. It is then obvious that Gyula Farkas turned to the research of the data of the house, the more so, because the finding and the identification of the birthplace had not quite been finished and beyond doubts. Thereby the fact, yet not mentioned even by Bolyai researchers, that Gyula Farkas collected 31 documents (more exactly case items) about the parental house of Janos Bolyai. Our aim is to examine these documents and to express our gratitude to Gyula Farkas, the Bolyai researcher.

AMS 2000 subject classifications: 01A70
Key words and phrases: Gyula Farkas, János Bolyai

All the people who ever looked more deeply into the work and writings of the two Bolyais were impressed by their greatness and could no longer be indifferent about them. The same thing it may have happened to Gyula Farkas, the most quoted professor of the former University of Kolozsvár. Gyula Farkas' greatness was also appreciated only by posterity. Today the impact factor and the introduction of international quotations evidently show that he was one of the most prominent professors of the University of Kolozsvár.

The confession of Gyula Farkas about the Bolyais

Let us remember the words of Gyula Farkas, uttered at thhe inauguration of the Bolyai memorial plaque, then in Marosvásárhely, at the reinterment of the Bolyais.

* email: olahgal@nextra.ro
** Partially supported by the Arany János Közalapítvány, Budapest

Fig. 1: Birth-house of János Bolyai.

Fig. 2: Present grave of two Bolyais.

There Dr. Gyula Farkas, Dean of The Faculty of Mathematics and Natural Sciences addressed the following speech to Mr. Antal Salamon city counsellor and deputy of Kolozsvár City Council:

Respected Sir of the Council!

The Faculty of Mathematics and Natural Sciences of The Ferencz József Hungarian Royal University of Sciences decided already three years ago to search for János Bolyaiś birthplace and to mark it with a memorial plaque. In the research it has been supported by the respected Council of our noble town.

The first traces led to another house that was owned by the landholder Bálint Betegh at the opposite side of Ferenc Deák street. At the time of János Bolyai s birth it was the house of Pál Bodor, the provincial controller of the Transylvanian Cashbox.

However, one of my professor colleagues found out that in this house our János Bolyai was born. It is presently owned by the favoured local Trade Association and at that time it had been the property of János Bolyai's maternal grandfather, József Benkő.

Having learnt this fact, The Faculty immediately gave effect to the setting up of the memorial by the very kind leave of the Trade Association. On the plaque János Bolyai is said to be the Hungarian Euclide, since he had been a creative master of geometry like Euclide. His father, Farkas Bolyai is also remembered on the plaque, he deserves this as the deep thinking author of the Tentamen, and also for supporting the development of János's mathematical talent.

I am turning to the respected Sir of The Council with the solid conviction that the worthy public of Kolozsvár Free Royal City will treasure the memory of

János Bolyai's natal house and I ask that the monument be kindly accepted in the protection of the respected city authorities.

Fig. 3: Original grave of János Bolyai.

Fig. 4: Present location of original grave.

With zealous words Antal Salamon accepted the monument in the protection of the town, promising to honor it in the greatest grace and goodwill.

Now the wreaths of the Hungarian Academy of Sciences, the Budapest Royal Hungarian University of Sciences, The Association of Mathematics and Physics, the public of Kolozsvár's Free Royal City and the The Ferencz József Hungarian Royal University of Sciences have been placed on the memorial, the text of which is the following:

In the year 1802, in the 12th month, on the 15th day, here was born János Bolyai Bolyai, the Hungarian Euclide, son of Farkas Bolyai Bolyai, who was the deep thinking author of the Tentamen. After 100 years of János Bolyai's birth The Faculty of Mathematics and Natural Sciences of The József Ferencz University of Sciences has paid this tribute. (1903)

The speech of corresponding member Gyula Farkas at the interment of János Bolyai's ashes next to his father's in Marosvásárhely, 1911, July 7.

The Bolyais!

A father searching into depths and a son seeing into distant spaces! On behalf of The Hungarian Academy of Sciences I salute your ashes. Behold, a careful organization brought your ashes together, since you too had been victims of separation in vain: you belong together not only by the material order of nature but also by the order of The one who rules over the countries of spiritual life until infinite times.

Since you have parted into the land of eternal peace, the common rooted and mingled branched trees of arithmetics and geometry have given fruits with plenty and ripened your theories as well and "The Appendix, Scientiam Spatii absolute veram exhibens" has come to great glory, and today the enormous natural law raises it to itself, the law of laws, calling our thinking into such a world of space and time where the Bolyai-world of space finds a new impulse.

In the light of your glory our Academy paid a tribute famous all around the world through the Bolyai Foundation, in your memory, who made the fame of the Hungarian state so much greater.

Alas, you had to be among those craftsmen of science who cannot live the triumph of their thoughts. But the bitterness that presses upon human frailty is very small compared to the huge recompense that rewards the best workers of cognition with cognition itself.

That is the reason that lessens our – your distant successors' – pain, be expressing our appreciation and gratefulness. In this manner I lay on your dear monuments the wreaths of the Hungarian Academy of Sciences. (1911)

These few sentences truly demonstrate the spiritual greatness of Gyula Farkas. Zoltán Gábos, academician, mathematician and physicist, the best specialist in Gyula Farkas' life and work, successor of Gyula Farkas' Department correctly remarked that "we are much indebted to Gyula Farkas." The presented documents prove that Gyula Farkas searched for the documents available on the Bolyais in Kolozsvár with great diligence and modesty. He bought these documents, made them a collection and found for them the best place possible, the Archives of the MTA's library. (The Library of the Hungarian Academy of Sciences).

This fact can be appreciated only by those who know that in the State Archives in Cluj the heritage of the professors of the Franz Joseph University of Sciences is still not available for research. (All inquiries are rejected with the excuse that these works are not yet arranged and registred.)

Gyula Farkas collected 31 documents about János Bolyai, and it is our responsability to transform these a common treasure by processing them.

References

[1] Biblioteca Teleki-Bolyai Târgu-Mureş, Romania. Bolyai János születésének századik évfordulója alkalmából a Kolozsvári M. Kir. Ferencz József Tudományegyetem által 1903 januárius 15-ikén rendezett emlékünnep. Kolozsvár 1903. Ajtai K. Albert Könyvnyomdája, 71-72.p.

[2] Library of the Hungarian Academy of Sciences, Department of Manuscript and Rare Books, Collection Bolyai.

Gyula Farkas and the Mechanical Principle of Fourier

Ferenc SZENKOVITS* **

Department of Applied Mathematics,
Babeş-Bolyai University Cluj-Napoca
Str. Cireşilor 19, RO-400487 Cluj-Napoca, Romania

Department of Astronomy, Eötvös Loránd University,
Pázmány Péter sétány 1/A, 1117 Budapest, Hungary

Abstract. Gyula Farkas (1847–1930), the great Hungarian scientist in mathematics and physics, professor of the University of Cluj (1887–1915) dedicated an important part of his scientific researches to the foundations of the mechanics. He focused mainly on the conditions of the mechanical equilibrium, dialing a more general form of the principle of the virtual work, the inequality form, known as the mechanical principle of Fourier, dedicating many scientific papers to this principle and to the related mathematical problems. Farkas dealt with the foundation of force equilibrium in mechanics and thermodynamics when he created the famous theorem of homogeneous linear inequality systems.

Gyula Farkas made not only researches dedicated to the principles of mechanics, but in his courses of Analytical mechanics he used the Fourier-principle for axiomatic foundations of the mechanics and in different applications.

In this paper a comprehensive presentation of Gyula Farkas' work dedicated to the mechanical Fourier-principle is given.

AMS 2000 subject classifications: 70-03,70A05, 90C90

Key words and phrases: Gyula Farkas, the mechanical principle of Fourier, history of mathematics

1 Introduction

It is widely accepted that Gyula Farkas (1847–1930) was one of the greatest Hungarian scientists in mathematics and physics. To the present days he was one of the most famous professor of the University of Cluj (Kolozsvár, Hungary at that time), where he worked twenty-eight years (1887–1915). The main part of his scientific activity is related to this years spent in Kolozsvár, where he made fruitful researches in various fields of the theoretical physics and mathematics.

In December 1892 Gy. Farkas participated in a celebration in Padova where the 300$^{\text{th}}$ year anniversary of Galilei's starting his activity there was held. One month later, in January 19, 1893 he presented the use of the virtual work in the papers of Galilei at the Mathematical and Physical Society in Budapest.

Starting from this year (1892) Gyula Farkas dedicated an important part of his scientific researches to the foundations of the mechanics. He focused on the

* email: fszenko@math.ubbcluj.ro

** Partially supported by the Hungarian Academy of Science, trough the János Bolyai grant: BO/00003/04.

conditions of the mechanical equilibrium, dealing with a more general form of the principle of the virtual work. He studied the inequality form of this principle, the so called Fourier-principle of the mechanics. In the next years he dedicated 9 scientific papers to this principle, and other 9 papers related to its mathematical problems (see [22]). Farkas dealt with the foundation of force equilibrium in mechanics and thermodynamics when creating the famous theorem of homogeneous linear inequality systems, the well known and so frequently cited *Farkas Theorem* (see [20] and [21]).

Gyula Farkas made not only researches dedicated to the principles of mechanics, but in his courses of Analytical mechanics presented at the Franz Josef University of Cluj he used the Fourier-principle for theoretical foundations of the mechanics and in different applications.

Our purpose is to make a comprehensive presentation of Gyula Farkas' work dedicated to the mechanical Fourier-principle.

2 Celebration of Galilei in Padova

In December 1892, Gyula Farkas representing the Franz Josef University of Cluj – as dean of the Faculty of Mathematics and Natural Science at that time – participated in a celebration in Padova where the 300th year anniversary of Galilei's starting his activity there was held. At this occasion the rector of the University of Padova conferred the degree of *doctor honoris causa* to Gyula Farkas (together with other representatives of different universities).

For this special anniversary Farkas made a solid preparation, with a deep study of Galilei's work. Coming back, he presented two talks in January 19, 1893 at the current session of the Mathematical and Physical Society (Math. Phys. Társulat) in Budapest. In the first one, entitled *About Galilei and the celebration of Galilei in Padova,* he presented this celebration [1]. In the second talk Farkas presented a study dedicated to the development of the principle of the virtual velocities in Galilei's works [2]. In this second talk he gave a detailed presentation of Galile's scientific papers dedicated to the problem of equilibrium. He described four fundamental works, with Hungarian translation of the important parts of these one. The short title of these works are: *Della Scienza Meccanica; Discorso intorno i Galleggeanti; Dialogo dei Massimi Sistemi; Dialoghi delle nuove Scienze.*

Farkas made a profound study of these works. We can see this reading his analysis about some inaccuracy in Galilei's work. He presented e.g. the next problem: Two bodies C and D with equal masses and total weight π are connected with an inextensible thread trough the fixed points A and B. The stretched thread is horizontal. A third mass H of weight p is attached to the center E of the thread (Fig. 1).

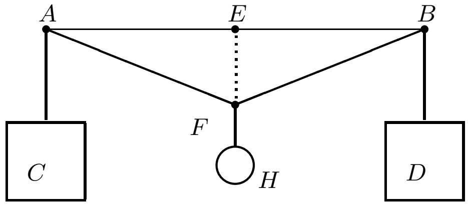


Fig. 1: Galilei's problem

Related to this problem *Galilei* asserts in his paper that the condition of the motion of the weight p downwards is:

$$\frac{z}{\zeta} > \frac{\pi}{p}, \tag{1}$$

where $z = EF$ is the vertical displacement of H and ς is the vertical displacement of C and D.

From the principle of the virtual displacements, Farkas deduce that the condition of the motion of the weight p downwards is $p\delta z > \pi\delta\zeta$. By using the conditions imposed trough the constraints, for the virtual displacements he found $z\delta z = r\delta r$, $\delta\zeta = \delta r$, $r = AF = BF$. In conclusion the condition of the motion is

$$\frac{r}{z} > \frac{\pi}{p} \tag{2}$$

which is the general correct form and differs from the condition (1) of Galilei.

Farkas also proved that the condition (1) is true only in the case when $z \to 0$.

It seems that this was the moment when the attention of Gyula Farkas was definitively captured by the principles of mechanics.

3 Farkas' first paper dedicated to the Fourier-principle

In 1894 Gy. Farkas publishes in Hungarian his first paper dedicated to the mechanical Fourier-principle. This paper entitled *Applications of the mechanical principle of Fourier* [3], was published in the same year in German too in *Math. und Nat. Berichte aus Ungarn* (*M.N.B.U.*).

The main purpose of the author in this paper is to show that *"the method of Lagrange's multipliers can be applied also, with small modifications, in the case of the Fourier-principle"*. Farkas underlines that this presentation is more general than the one given by Ostrogradsky (1834), because *Ostogradsky* imposed the condition that *the number of relations expressing the constraints to be not greater then those of the virtual displacements.*

Farkas used the next formulation of the problem:

If the mass points with masses $m', m'', \ldots$ with Cartesian coordinates $p', p'', \ldots$ and with the virtual changes of these coordinates $\delta p', \delta p'', \ldots$ verify the constraining conditions

$$\sum A\delta p = 0, \qquad \sum B\delta p = 0, \ldots \tag{3}$$

$$\sum L\delta p \geq 0, \qquad \sum M\delta p \geq 0, \ldots \tag{4}$$

then the condition of the equilibrium (the Fourier-principle) is:

$$\sum P\delta p \leq 0, \tag{5}$$

where P is the free force acting on the coordinate p; in case of the motion the d'Alembert's principle states that

$$\sum (P - m\ddot{p})\,\delta p \leq 0, \tag{6}$$

where $\ddot{p}$ denotes the acceleration in the variation of the coordinate p. All these sums are made for all the coordinates.

Farkas makes the remark that this general principle of Fourier, which contains inequalities, is more general than the classical principle of the virtual work (also called principle of the virtual velocities) expressed only with equations. This is evident if we see that in case of the constraints expressed only by equations, for any system of virtual displacements (δp) the opposite $(-\delta p)$ is also a virtual displacement and from $\sum P\delta p \leq 0$ and $\sum P\,(-\delta p) \leq 0$ we have $\sum P\delta p = 0$.

In the first part of this paper, following the introduction, entitled *Algebraic preliminaries*, Farkas presents a first formulation and (not complete, [16], [20], [21]) demonstration of his famous theorem:

Theorem 1 (Farkas). *If any solution $u, v, \ldots$ of the system*

$$\left.\begin{array}{c} A_1 u + B_1 v + \ldots \geq 0, \\ A_2 u + B_1 v + \ldots \geq 0, \\ \textit{etc.} \end{array}\right\} \tag{7}$$

is also a solution of the inequality

$$A_0 u + B_0 v + \ldots \geq 0, \tag{8}$$

then exist the positive coefficients $\lambda_1, \lambda_2, \ldots$ for which we have

$$A_0 = \sum \lambda_i A_i, \quad B_0 = \sum \lambda_i B_i, \ldots$$

In the second part: *The basic method of the application,* Farkas presents the possibility of the application of the method of Lagrang's multipliers in the case of the Fourier-principle, to eliminate the virtual displacements and to deduce the so called Lagrange's equation of the first kind of the mechanics.

Farkas indicates that in the first step is recommended to reduce the number of the independent coordinates and virtual displacements, expressing the variations δp from the constraints (3) by using other independent virtual displacements δq. The remaining equations and inequalities expressing the constraints will be:

$$\sum F\delta q = 0, \qquad \sum G\delta q = 0, \ldots \tag{9}$$

$$\sum S\delta q \geq 0, \qquad \sum T\delta q \geq 0, \ldots \tag{10}$$

and the inequalities expressing the principle – deduced from (5) or (6) in case of equilibrium, respectively the motion – will be:

$$\sum Q\delta q \leq 0 \qquad \Leftrightarrow \qquad -\sum Q\delta q \geq 0. \tag{11}$$

and

$$\sum (Q - m\ddot{q})\, \delta q \leq 0. \tag{12}$$

Equations (9) will be expressed also by using inequalities, and the system of the constraints will be:

$$\begin{aligned} \sum F\delta q \geq 0, \qquad & \sum G\delta q \geq 0, \ldots \\ -\sum F\delta q \geq 0, \qquad & -\sum G\delta q \geq 0, \ldots \\ \sum S\delta q \geq 0, \qquad & \sum T\delta q \geq 0, \ldots \end{aligned} \tag{13}$$

The mechanical principle of Fourier requires that any solutions δq of (13) to be also solutions for (11). By using now the Farkas theorem results that there exists some positive multipliers which allows us to express the coefficients $-Q$ as linear combinations of the coefficients $F, G, \ldots, -F, -G, \ldots$and S, T. Let denote these coefficients by $\varphi', \psi', \ldots, \varphi'', \psi'', \ldots$. and λ, μ. Corresponding to the Farkas theorem

$$-Q = \left(\varphi' - \varphi''\right) F + \left(\psi' - \psi''\right) G + \ldots + \lambda S + \mu T + \ldots$$

But the differences $\varphi' - \varphi''$, $\psi' - \psi''$,... may be positive or negative to. In conclusion, the Fourier-principle is verified for coefficients Q with

$$Q + \varphi F + \psi G + \ldots + \lambda S + \mu T + \ldots = 0,$$

where $\varphi, \psi, \ldots$ are arbitrary real parameters, and $\lambda, \mu, \ldots$ not negative arbitrary parameters.

This first paper contains other two parts dedicated to *Auxiliary methods* and to the *Two mine types of the applications*, where he presents *Mechanical equations of rigid bodies in contact* and *The general mechanical equations of not rigid bodies.*

4 Further works about the Fourier-principle

In 1895 Farkas published a series of three papers dedicated to the Fourier-principle under the title *The history of the Fourier-principle, and some special applications I-III [4]*. In the first part – *The history of the principle* – he presents the first formulation of the principle in the inequality form by Fourier (1798), the apparition of the principle in the works of Gauss (1829). Important contributions to the development of this principle were made by Ostogradsky (1834), who expressed his astonishment on the fact that Lagrange, in the new edition of the famous *Mécanique analitique* (1811) did not used this principle and established the possibility of the application of the method of Lagrange's multiplicators, but only in the case, when tho number of restricting inequalities is less than the number of virtual displacements.

Farkas mentions also the contemporary works of Rausenberger (*Lehrbuch der analytischen Mechanic*, 1888) and Schell (*Theorie der Bewegung und der Kräfte*) which contains short presentations of the principle.

In conclusion Farkas sadly write that the Fourier-principle is almost forgotten, used only by a few authors in some particular cases, and the general case was studied only by Ostogradsky. He underline also that the possibility of the application of the general form was neglected.

In the next two parts of this paper he gives variate types of concrete applications, such as: the equilibrium of the mass point on a resisting surface, the equilibrium of the mass point at the intersection of two resisting surfaces, the equilibrium of the mass point at the intersection of three resisting surfaces, ...,the motion of the mass point with friction on a resisting surface. The last part contains 6 examples of different cases of equilibrium and motion of the rigid body with unilateral restrictions.

In the following years Farkas published other five papers dedicated mainly to the algebraic foundation of the applications of the mechanical principle of Fourier [5], [6], [7], [10], [11]. In these papers he these papers he gave more and more elaborated and complete demonstration of his famous theorem.

The mechanical principle of Fourier is present also in the lectures given by Farkas at the Franz Josef University in Cluj. We can see this reading Farkas' handbooks multiplied for students: *Theory of vectors and simple inequalities* [8], *Foundations of mechanics* [12], *Analytical mechanics* [13], containing profound, detailed, accurate presentation of the mathematical basis of the Fourier principle, the foundation of the analytical mechanics based on this principle and variate concrete applications of this principle.

5 Conclusions

After this review of Farkas' works dedicated to the mechanical principle of Fourier, and taking account also on the appreciations made by Fényes [15], Filep [16], Prékopa [20] and [21], Martinás [19], Gábos [17] and Rapcsák [22], we can conclude, that Gyula Farkas work dedicated to the mechanical principle of Fourier are very important especially for the further development of the optimization theory, where the Farkas theorem is a central result.

Unfortunately at present days the mechanical principle of Fourier is almost forgotten. Basic books dedicated to the foundations of the mechanics do not use this general principle, they use only the particular case of the bilateral constraints, which can be expressed only by equations. In the best case the problem of unilateral constraints is only mentioned as in Gantmacher [18]: "The motion of a system on which a unilateral constraint is imposed may be divided into portions. . . , in certain portions a unilateral constraint is either replaced by a bilateral constraint or is eliminated altogether. For this reason we shall henceforth consider only bilateral constraints."

References

[1] Farkas, Gyula, Galileiről s a páduai Galilei-ünneplésről. *Math. és Phys. Lapok* **2** (1893), 65–77. [About Galilei and the celebration of Galilei in Paova – in Hungarian]

[2] Farkas, Gyula, A virtuális sebességek elve Galileinél. *Math. és Phys. Lapok* **2** (1893), 78–95. [The principle of the virtual velocities at Galilei – in Hungarian]

[3] Farkas, Gyula, A Fourier-féle mechanikai elv alkalmazásai. *Mat. és Termtud. Értesítő* **12** (1894), 457–472. The same in German: *Math. und Nat. Berichte aus Ungarn* (*M.N.B.U.* in the next) **12** (1894), 263–281. [Applications of the mechanical Fourier-principle – in Hungarian]

[4] Farkas, Gyula, A Fourier-féle mechanikai elv története és némely speciális alkalmazása I–II–III. *EME Értesítője* **20** (1895), 43–54, 212–221, 326–335. [History and some special applications of the mechanical principle of Fourier – in Hungarian]

[5] Farkas, Gyula, A Fourier-féle mechanikai elv alkalmazásának algebrai alapjáról. *Math. és Phys. Lapok* **5** (1896), 49–54. The same in German: *M.N.B.U.* **16** (1899), 154–157. [About the algebraical fundaments of the application of the mechanical Fourier-principle – in Hungarian]

[6] Farkas, Gyula, A Fourier-féle mechanikai elv alkalmazásainak algebrai alapjai. *Mat. és Termtud. Értesítő* **16** (1898), 361–364. The same in German: *M.N.B.U.* **16** (1899), 25–40. [Algebraical fundaments of the application of the mechanical Fourier-principle – in Hungarian]

[7] Farkas, Gyula, Paraméteres módszer Fourier mechanikai elvéhez. *Math, és Phys. Lapok* **7** (1898), 63–71. [Parametric method for the mechanical Fourier principle – in Hungarian]

[8] Farkas, Gyula, *Vektor-tan és az egyszerű inequatiok tana.* Tudományegyetemi előadásaiból közli Dr Farkas Gyula a M. Tud. Akadémia L. tagja.

Különlenyomat az Erdélyi Múzeum-Egylet Orvos- és Természettudományi Szakosztályának Értesítője után. Kolozsvár, Stein János M. Kir. Tud. Egyetemi Könyv-kereskedése, 1900. [Theory of vectors and simple inequalities – in Hungarian]

[9] FARKAS, GYULA, Általános mechnikai elvek az aether számára *Mat. és Termtud. Értesítő* **19** (1901). The same in: *"Archiv Nederlandaises Livre jubilaire dédié à H. A. Lorentz"* 1900, 56—75. [General mechanical principles for the aether – in Hungarian]

[10] FARKAS, GYULA, Theorie der einfachen Ungleichungen. *Journal für die Reine und Angewandte Mathematik (Crelle's Journal)*, **124** (1902) 1–27.

[11] FARKAS, GYULA, Beiträge zu der Grundlagen der analytischen Mechanic. *Journal für die Reine und Angewandte Mathematik (Crelle's Journal)* **131** (1906) 165–201.

[12] FARKAS, GYULA, *A mechanika alaptanai.* Sokszorosított egyetemi jegyzet: M. Kir. Ferenc József Tudományegyetem, Kolozsvár 1914. (Előadta: Dr. Farkas Gyula e. ny. r. tanár, 1913–1914 tanév II. felében.) [Foundations of mechanics – in Hungarian]

[13] FARKAS, GYULA, *Analitikus mechanika.* Sokszorosított egyetemi jegyzet: M. Kir. Ferenc József Tudományegyetem, Kolozsvár 1908, 1914. (Előadta: Dr. Farkas Gyula e. ny. r. tanár. 1913–14 tanév I. felében.) [Analytical mechanics – in Hungarian]

[14] FARKAS, GYULA, *Erőtan.* Sokszorosított egyetemi jegyzet: M. Kir. Ferenc József Tudományegyetem, Kolozsvár 1914. (Előadta: Dr. Farkas Gyula e. ny. r. tanár, 1913–14 tanév II. felében.) [Theory of forces – in Hungarian]

[15] FÉNYES, IMRE, Megjegyzések és kiegészítések a mechanika elveinek Farkas Gyula-féle tárgyalásmódjához. *Fizikai Szemle* **4** (1954) 99–102. [Observations and complements to the Gyula Farkas' discussions of the principles of the mechanics – in Hungarian]

[16] FILEP, LÁSZLÓ, Farkas Gyula élete és munkássága. *Matematikai Lapok* **29** (1977–81) 231–244. [Life and work of Gyula Farkas – in Hungarian]

[17] GÁBOS, ZOLTÁN, "A természet a matematika nyelvén szól hozzánk" *Természet Világa* **128** (1997) 290-293. ["The nature speaks us on the language of the mathematics" – in Hungarian]

[18] GANTMACHER, F., *Lectures in Analitical mechanics*, Mir Publishers, Moscow, 1975.

[19] MARTINÁS, KATALIN, Farkas Gyula (1847–1930), *Fizikai Szemle* **8** (1992) 298–301.

[20] PRÉKOPA, ANDRÁS, Az optimizáláselmélet kialakulásának történetéről, *Alkalmazott Matematikai Lapok* **4** (1978) 165–161. [On the development of optimization theory – in Hungarian]

[21] PRÉKOPA, ANDRÁS, On the development of optimization theory, *American Mathematical Monthly* **87** (1980) 527-542.

[22] Rapcsák, Tamás, Mechanical equilibrium and equilibrium systems. *Equilibrium problems and variational models*, P. Daniele, F. Giannessi and A. Maugeri eds., Kluwer Academic publisher, 2003, 379–399.

Computer Assisted Teaching of Operations Research at the University of Kaposvár

Eleonóra STETTNER* and Gyöngyi BÁNKUTI**

University of Kaposvár
Institute of Information Technology
Department of Applied Mathematics and Physics

Abstract. Our Operations Research program to be presented has been developed for MEng students of agriculture and agricultural economy (general course and specialization in information technology), as well as for MSc students of business administration. The main topics are the linear and nonlinear programming, including parametric, hyperbolic, integer, quadratic, etc. methods. The lectures are followed by practical exercises held in computer labs.

We intended to give the students a tool that helps in the routine calculations but demonstrates the applied algorithms transparently. However the available computer assisted teaching tools either provide complete solutions, without the direct access of the students to the detailed insight, or the actual methods are embedded in large mathematical packages like MathLab.

Considering this we concluded to develop a special software tool for above described field of applications. In addition this package helps in the teaching of linear algebra, as well. The basic capabilities of the tool cover matrix multiplication, matrix inverting, solving systems of linear equations, as well as solution of complete linear, hyperbolic, parametric programming problems. The elaboration of the software was supported by a regional grant.

The students use the package as an auxiliary toolkit, and, with the knowledge of the problem to be solved, they are forced to decide about the subsequent steps of the Simplex Method, while the procedure of solving is accelerated by the computer aided execution of the individual routine calculations.

AMS 2000 subject classifications: 90-01

Key words and phrases: operations research, teaching

At the University of Kaposvár we teach Operations Research for general agricultural engineers and agricultural engineers specialized in economics. The major topics in the first course are the followings:

1. Vector, linear space, (independence, basis transformation etc.)
2. Matrix, definition, operations, inverse,
3. Linear programming, simplex method, (graphical representation)
4. Special cases: degeneration, alternative optimum
5. Duality-sensitivity analysis
6. Hyperbolical goal functions
7. Binary, integer and mixed binary-integer programming
8. Parametrical programming
9. Transportation problem

* email: stettnere@freemail.hu
** email: bankuti@matinf.gtk.u-kaposvar.hu

10. Project planning (CPM, PERT)
11. Simple business model of macro-economical connections
12. Modeling, simulation

We have also an advanced course for the students specialized in informatics (while it can also be taken by the other students optionally, as well.) The topics in this course are the followings:

1. Nonlinear programming: quadratic (KKT constraints)
2. Multiple goal functions: sequential optimization, weight functions, dimensionless functions,
3. Efficient point: definition, iterative method for production
4. Dynamic programming,
5. Basics of two person zero sum game theory
6. Stochastic processes, Markov chains
7. Queuing theory
8. Bayesian decision models
9. Inventory control: deterministic models, stochastic models, discount models
10. Probability theory
11. Data mining (Bayesian information criteria)
12. Artificial Intelligence, search methods

In this course the advanced module of operational research is supplemented with some other related fields (like artificial intelligence, data mining, etc.), because the students do not learn these special topics and tools in other courses.

Considering that nowadays computer applications have keynote role, we involve three generally available or simple software tools in the curriculum, as follows:

– Excel Solver for solving models of linear programming,
– MS Project for project planning, and
– our own educational software for the linear and some nonlinear programming methods to support the manual calculation carried out by the Simplex method)

In this paper mostly the use of the latest computer assisted education methodology is introduced.

The **Linv** program can be applied for the solution of the following problems:

– matrix multiplication,
– determination of the inverse of a matrix,
– solution of linear equations,
– solution of linear programming problems,
– solution of multi-objective LP problems,
– solution of linear integer programming problems,
– solution of parametric optimization problems,
– solution of hyperbolic optimization problems.

The pedagogical philosophy of the applied methodology is that the computer tool must not take the responsibility for the feasibility of data and for the correct procedure of problem solving, only helps to execute the slow manual computational steps. Consequently the student can (and has to) concentrate on the essential algorithm and on the procedure of problem solving, itself. The investigated problem can be saved and reloaded at various stages, as well.

The main task window can be seen in Fig. 1.

Fig. 1. The main screen of the Linv program

The first option is the matrix multiplication, where the arrangement of the matrices is the same as it used in teaching: The multiplicand matrix is on the left side of the sheet, while the multiplier matrix is set at the upper side and the result appears in the middle. The dimensions of the multiplicand and the number of the columns of the multiplier can be specified in the dialog windows. We can choose either decimal or fractional representation of numbers. The program calculates immediately when the inlet data are specified, while the result matrix shows the actual state of multiplication to the student (See Fig. 2).

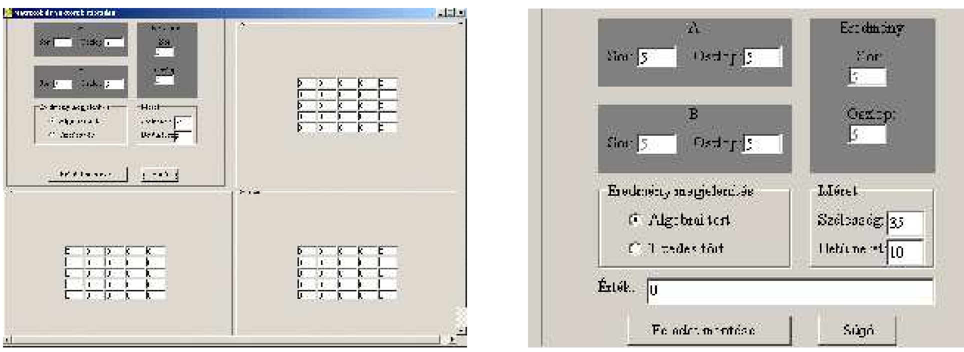


Fig. 2. Matrix multiplication screen

In the matrix inversion (see Fig. 3), we can select between the decimal and fractional representation, too. After having filled the matrix (by the mouse and/or tab key) the "pivot element" can be designated by double mouse click. Then the calculation of a single step of the simplex method becomes possible. Of course, the students have to know how to the inverse of a matrix is applied in the simplex method. The program saves the consecutive states of the tables, and it is possible to go backward and forward stepwise between the first and the last steps.

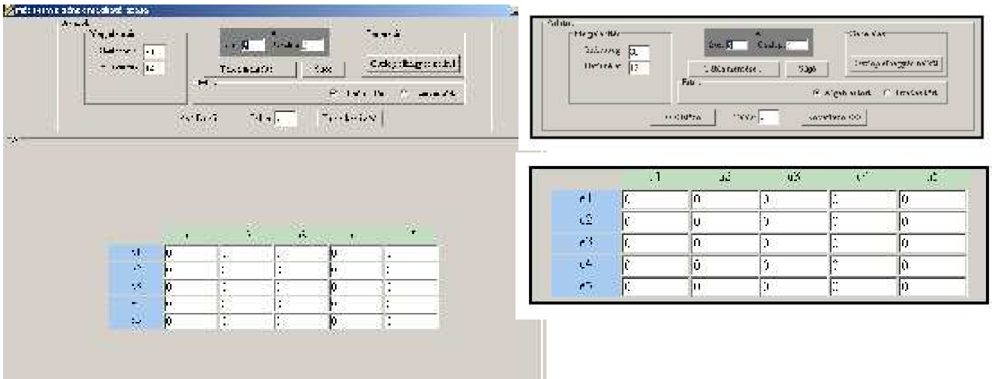


Fig. 3. Matrix inverting screen of the program

The Linear Programming Module and the Linear Equation Solver works with simplex method, as well. Having filled the trivial representation of the problem, we can also designate by double mouse click the pivot element, and then the program calculates the individual steps of the simplex method.

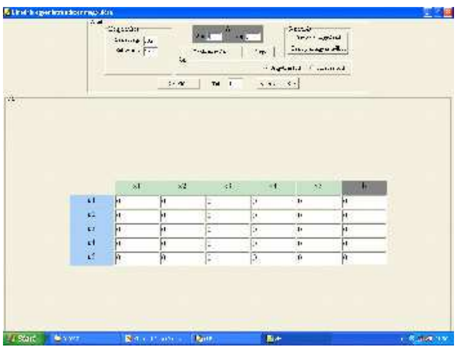


Fig. 4. Screen of the Linear Equation Solver

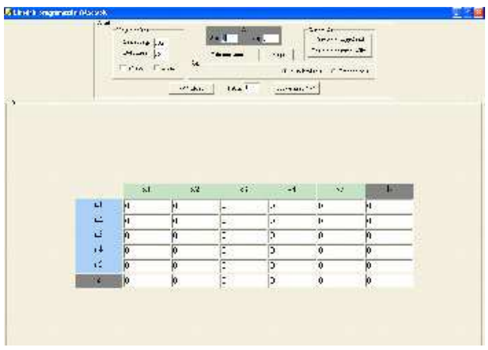


Fig. 5. Screen of the LP Solver

The additional options in this task are the followings:

- Adding a "secondary goal function" row (z^*) if it is necessary;

• Choosing dual representation and modifying the names of the variables (u_i, x_i, z) for u_i* , and $-u_i$, or for anything else in the respective dialog box;

• We can hide the pivot elements column if it is needed. We used to hide it in the two stage simplex method.

The non-linear problems are similarly represented as they used to be in the manual calculation. To the solution of these problems the same simplex method is used with the same options.

In the multi-objective LP problem the additional option of the program is the consideration of the previously defined coefficients of the goal functions. While solving parametric programming problems the computer program manipulates also the additional rows for the constants and of the parameters (z_c, z_t). In the hyperbolic case, new rows for the coefficients of the numerator and denominator functions are added. In the integer programming menu, new rows, Gomory Cutting Planes can be added.

A	B	C	D	E	F	G	H	I
1		x_1	x_2	x_3				
2	u_1	1	2	1	4			30
3	u_2	1	0	1	2			20
4	u_3	2	1	1	4			35
5								
6	„-z"	2	1	4	7			0
7								
8		1	1	1				

```
"G2"=SUM & MULTIPLICATION(B2:D2;B$8:D$8)
"G3"=SUM & MULTIPLICATION(B3:D3;B$8:D$8)
"G4"=SUM & MULTIPLICATION(B4:D4;B$8:D$8)
"G6"=SUM & MULTIPLICATION(B6:D6;B$8:D$8)
```

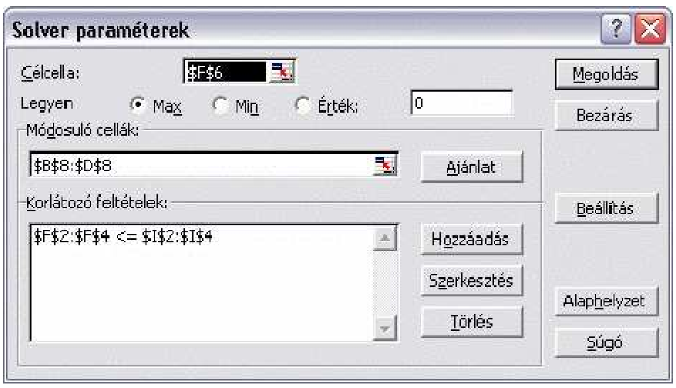


Fig. 6. Part of the brief manual of Excel Solver for LP problems

Instead of the unnecessary manual, a detailed help supports the use of our simple program.

The second computational tool used during the instructions is Excel Solver. A brief booklet describing this tool is issued, just to help the less skilled students. In what follows we show only some parts of this textbook. (Fig. 6.)

Excell Solver provides even sensitivity analysis besides the optimal solution. Students are taught how to interpret it in an economic environment.

Another booklet is provided on solving transportation problems by Excel Solver.

Third computational tool used by us for project planning is MS Project. We connect the manual PERT methodology (Fig. 7.) with the MS Project representation, and show the students the facilities of this problem (Critical Path visualization, Tracking GANNT, Resource Sheet, Resource Graf, PERT Chart etc.). (Fig. 8.)

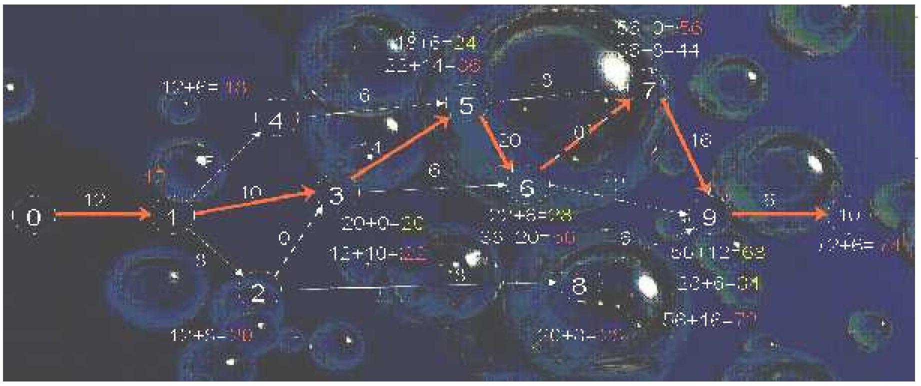
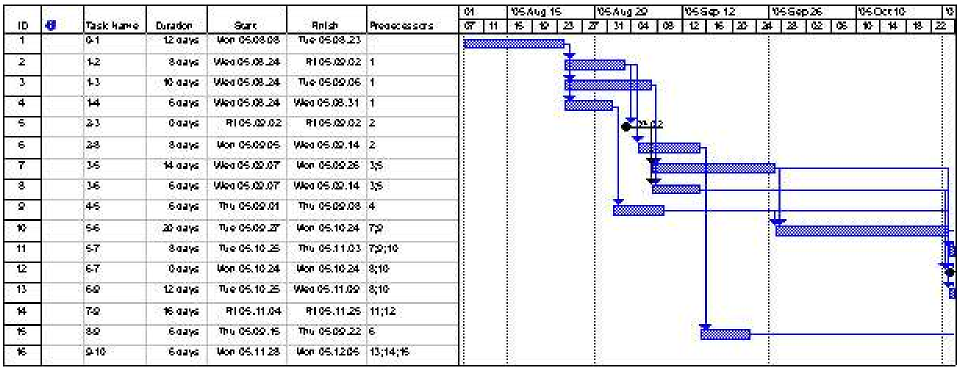


Fig. 7. . Project Planning, PERT Chart View (Event Graph)



ID		Task Name	Duration	Start	Finish	Predecessors
1		0-1	12 days	Mon 05.08.08	Tue 05.08.23	
2		1-2	8 days	Wed 05.08.24	Fri 05.09.02	1
3		1-3	10 days	Wed 05.08.24	Tue 05.09.06	1
4		1-4	6 days	Wed 05.08.24	Wed 05.08.31	1
5		2-3	0 days	Fri 05.09.02	Fri 05.09.02	2
6		2-8	8 days	Mon 05.09.05	Wed 05.09.14	2
7		3-5	14 days	Wed 05.09.07	Mon 05.09.26	3;5
8		3-6	6 days	Wed 05.09.07	Wed 05.09.14	3;5
9		4-5	6 days	Thu 05.09.01	Thu 05.09.08	4
10		5-6	20 days	Tue 05.09.27	Mon 05.10.24	7;9
11		5-7	8 days	Tue 05.10.25	Thu 05.11.03	7;9;10
12		6-7	0 days	Mon 05.10.24	Mon 05.10.24	8;10
13		6-9	12 days	Tue 05.10.25	Wed 05.11.09	8;10
14		7-9	16 days	Fri 05.11.04	Fri 05.11.25	11;12
15		8-9	6 days	Thu 05.09.15	Thu 05.09.22	6
16		9-10	6 days	Mon 05.11.28	Mon 05.12.05	13;14;15

Fig. 8. Project Planning in MS Project (GANNT Chart view)

References

[1] Bánkuti, Gy., Paál, J., State of art of the written examination in teaching of up to date mathematical economics, Hungarian Meeting of Teaching Physics, Mathematics and Computer Science in Higher Education, Pécs, Sept 7–9, 1995, Proceedings p. 21.

[2] Bánkuti. Gy., Paál J., Development trends of fundamental curriculum subjects int he agricultural higher education., Hungarian Meeting of Teaching Physics, Mathematics and Computer Science in Higher Education, Pécs, Sept 7–9, 1995, Proceedings p. 26.

[3] Dr. Csernyák László, Dr. Jánosa András, *A gazdasági optimalizálás módszerei II. Lineáris optimalizálás*, Nemzeti Tankönyvkiadó, 2004.

[4] Hillier, F. S., Lieberman, G. J., *Bevezetés az operációkutatásba*, LSI Oktatóközpont, 1994.

[5] Stettner Eleonóra, Tevelinė Matejdesz Anikó, *Lineáris algebra elektronikus jegyzet a Kaposvári Egyetem gazdálkodás szakos hallgatói számára*, 2004.

[6] Stettner Eleonóra, *Operációkutatás hallgatói segédlet, kidolgozott feladatok, elektronikus formában*, 2004.

[7] Softwares: *MS Excel, Solver; MS Project; Linv Program*, Load from: `matinf.gtk.u-kaposvar.hu`

List of participants

Bánhelyi Balázs, Szeged, Hungary, banhelyi@inf.u-szeged.hu
Bánkuti Gyöngyi, Kaposvár, Hungary, bankuti@matinf.gtk.u-kaposvar.hu
Blaga, Petru, Cluj, Romania, blaga@math.ubbcluj.ro
Bódizs Tünde, Cluj, Romania, bodizst@.yahoo.com
Boţ, Radu Ioan, Chemnitz, Germany, bot@mathematik.tu-chemnitz.de
Csendes Tibor, Szeged, Hungary, csendes@inf.u-szeged.hu
Cseke Botond, Cluj, Romania csekebotond@yahoo.co.uk
Csizmadia Zsolt, Budapest, Hungary, csisza@cs.elte.hu
Egri Edit, Miercurea Ciuc, Romania, egriedit@email.ro
Fülöp János, Budapest, Hungary, fulop@oplab.sztaki.hu
Gaal György, Cluj, Romania
Gábos Zoltán, Cluj, Romania, zgabos@phys.ubbcluj.ro
Grad, Sorin-Mihai, Chemnitz, Germany, grad@mathematik.tu-chemnitz.de
Hodrea, Bogdan, Chemnitz, Germany, bogdan.hodrea@s2002.tu-chemnitz.de
Horváth Zoltán, Győr, Hungary, horvathz@sze.hu
Illés Tibor, Budapest, Hungary, illes@cs.elte.hu
Kao, Chiang, Tainan, Taiwan, R. of China, ckao@mail.ncku.edu.tw
Kassay Gábor, Cluj, Romania, kassay@math.ubbcluj.ro
Kása Zoltán, Cluj, Romania, kasa@cs.ubbcluj.ro
Kolumbán József, Cluj, Romania, kolumban@math.ubbcluj.ro
Komlósi Sándor, Pécs, Hungary, komlosi@ktk.pte.hu
Kovács Lehel István, Cluj, Romania, klehel@cs.ubbcluj.ro
Kuba Attila, Szeged, Hungary, kuba@inf.u-szeged.hu
Marchiş Julianna, Cluj, Romania, marchis_julianna@yahoo.com,
imarchis@math.ubbcluj.ro
Martinás Katalin, Budapest, Hungary, martinas@ludens.elte.hu
Moldovan, Andreea, Pisa, Italy, moldovan@mail.dm.unipi.it
Mordukhovich, Boris S., Detroit, USA, boris@math.wayne.edu
Mureşan, Marian, Cluj, Romania, mmarian@math.ubbcluj.ro
Nagy László, Cluj, Romania, lnagy@phys.ubbcluj.ro
Nagy Mariann, Budapest, Hungary, nmariann@cs.elte.hu
Oláh-Gál Róbert, Miercurea Ciuc, Romania, olahgal@nextra.ro
Popovici, Nicolae, Cluj, Romania, popovici@math.ubbcluj.ro
Prékopa András, Budapest, Hungary, prekopa@rutcor.rutgers.edu,
prekopa@cs.elte.hu
Révész Szilárd György, Budapest, Hungary, revesz@renyi.hu
Simon, Simion, Cluj, Romania, simon@phys.ubbcluj.ro
Stettner Eleonóra, Kaposvár, Hungary, stettner@matinf.gtk.u-kaposvar.hu
Szabó Péter Gábor, Szeged, Hungary, pszabo@inf.u-szeged.hu

Szenkovits Ferenc, Cluj, Romania, fszenko@math.ubbcluj.ro
Szilágyi Pál, Cluj, Romania, szilagyi_pal@sapientia.ro
Vizvári Béla, Budapest, Hungary, vizvari@cs.elte.hu
Wanka, Gert, Chemnitz, Germany, gert.wanka@mathematik.tu-chemnitz.de
Zălinescu, Constantin, Iaşi, Romania, zalinesc@uaic.ro

The talks of the conference

August 23^{rd}

Opening ceremony: L. Nagy, A. Prékopa, P. Blaga, S. Simon

A. Prékopa: Linear inequalities, duality theorem and their financial applications

G. Kassay, J. B. G. Frenk: Lagrangian duality and cone convexlike functions

C. Kao: Linear programming with interval parameters

R. I. Boţ, B. Hodrea, G. Wanka: A weaker reguralty condition for subdifferential calculus and Fenchel duality in infinite dimensional spaces

T. Illés: An elementary and constructive proof of the Farkas lemma

Zs. Csizmadia, T. Illés, M. Nagy, T. Terlaky: Linear complemenatrity problems, suffucient matrices and EP theorem

B. Filiz, Zs. Csizmadia, T. Illés: A new anti-degeneracy method for linear feasibility problems

August 24^{th}

Z. Gábos, L. Nagy: The investigation of Gyula Farkas in the field of electrodynamics and relativity theory

R. I. Boţ, G. Wanka: Farkas-type results with conjugate functions

B. S. Mordukhovich: Generalized diferentiation in variational analyses with applications in optimization

S. Komlósi: On quasiconvex farkas theorem

J. Kolumbán: Alternative inclusions

B. I. Boţ, S. M. Grad, G. Wanka: Conjugate duality for composed convex optimization problems with a new contraint qualification

A. Kuba: Farkas' lemma and its applications in discrete tomography

B. Bánhelyi, T. Csendes, L. Hatvani, B. Garay: A computer assisted proof and location of chaos: the case of a forced damped pendulum equation

Gy. Bánkuti, E. Stettner: Computer assisted teaching of Operations Research at the University of Kaposvár

J. Fülöp, S. Z. Németh: Global optimization techniques for stability analyses of decision functions

R. Oláh-Gál: Gyula Farkas as a Bolyai researcher

August 25^{th}

F. Szenkovits: Gyula Farkas and the Fourier principle
K. Martinás: Thermodynamic achievements of Gyula Farkas
C. Zălinescu: On the maximisation of (not neccessarily) convex function on convex sets
B. Vizvári: A little theory for the control of an assembly robot using Farkas theorem
N. Popovici: Lexicographic quasiconvex vector optimization
Z. Horváth: Invariant cones and polyhedra for dynamical systems
B. Farkas, Sz. Gy. Révész: Potential theory and randezvous numbers
M. Mureşan: Controlability and relaxation in Banach spaces