Zbornik radova 10(18)

TWO TOPICS IN MATHEMATICS

Matematički institut SANU

Zbornik radova 10(18)

TWO TOPICS IN MATHEMATICS

Editor:
Bogoljub Stanković

Matematički institut SANU

Издавач: Математички институт САНУ, Београд, Кнеза Михаила 35
серија: Зборник радова, књига 10(18)
За издавача: Богољуб Станковић, главни уредник
Уредник: Богољуб Станковић
Технички уредник: Драган Благојевић
Штампа: "Академска издања", Земун
Штампање завршено октобра 2004.

CIP – Каталогизација у публикацији
Народна библиотека Србије, Београд

519.176

TWO topics in Mathematics / editor Bogoljub Stanković.
– Beograd : Matematički institut SANU, 2004 (Zemun : Akademska izdanja).
– 91 str. : tabele ; 24 cm.
– (Zbornik radova / Matematički institut SANU ; 10(18))
Tiraž 350. – Str. 3: Preface / Bogoljub Stanković.
– Bibliografija uz svako poglavlje.

ISBN 86–80593–36–2

1. Stanković, Bogoljub
а) Теорија графова
COBISS.SR-ID 117351180

PREFACE

The aim of *Zbornik radova* is to foster further growth of pure and applied mathematics, publishing papers which contain new ideas and scopes in the mathematics. The papers have to be prepared in such a manner that they can inform readers in a favourable way, introducing them in a narrower field of mathematical theories pointing at research possibilities. It can be for the individual use or for discussions in College or University seminars.

We are open for contacts and cooperations.

Bogoljub Stanković
Editor-in-Chief

Contents

Dragoš Cvetković, Mirjana Čangalović
and Vera Kovačević-Vujčić

OPTIMIZATION AND HIGHLY INFORMATIVE GRAPH INVARIANTS

CONTENTS

ABSTRACT. It is known that graph invariants, which contain a great quantity of information on graph structure (for example, spectral invariants), are obtained by solving some extremal problems on graphs. Recently, such highly informative graph invariants are applied in solving optimization problems on graphs (e.g., the travelling salesman problem (TSP)). Using these paradigms, several relations, interconnections and interactions between graph theory and mathematical programming are described in this study. A model of TSP based on semidefinite programming and algebraic connectivity of graphs is described. A class of relaxations of this TSP model is defined and some solution techniques based on this class are proposed. Several examples of graph invariants defined by some kind of optimization tasks are also presented. Using several spectrally based graph invariants we treat the graph isomorphism problem.

1. Introduction

In this study we want to elaborate the following two assertions:

Assertion 1. Graph invariants, which contain a great quantity of information on graph structure, are obtained by solving some extremal problems on graphs.

Assertion 2. Highly informative graph invariants are useful in solving optimization problems on graphs.

If these assertions were mathematical statements, they should be proved in mathematical sense. We believe that they are true in an informal sense. Our experience in research shows much evidence of their validity. In this study we shall present several mathematical results which support them. Using these paradigms, several relations, interconnections and interactions between graph theory and mathematical programming are described.

In this introductory section we present some of the basic results from mathematical programming and graph theory which are necessary for the presentation of main ideas in Sections 2 and 3. In 1.1 an important optimization problem, the travelling salesman problem, is introduced. A highly informative graph invariant, the spectrum of a graph, is described in 1.2. Subsection 1.3 is devoted to semidefinite programming, a recently developed optimization technique and an important branch of mathematical programming.

Section 2 elaborates Assertion 2 while Section 3 elaborates Assertion 1.

Key words and phrases. Graph spectra, Algebraic connectivity, Graph isomorphism problem, Semidefinite programming, Travelling salesman problem, Branch-and-bound methods, Complexity indices, Clustering problems.

1.1. Travelling Salesman Problem. There is partly a joke, partly an advice in mathematics saying that if you do not know how to solve a problem you should find the first derivative and make it equal to zero. The point is that a great number of mathematical problems are optimization problems or can be reduced to them.

We shall begin with an exception.

Suppose that a salesman, starting from his home city, is to visit exactly once each city on a given list of cities and then to return home. It is reasonable for him to select the order in which he visits the cities so that the total of the distances travelled in his tour is as small as possible. This problem is called the *travelling salesman problem* (TSP).

TSP is a typical problem of *combinatorial optimization.* There is an extensive literature on and an impressive theory of TSP. The theory includes algorithms and heuristics (with an emphasis on complexity questions) for solving TSP as well as several variations and related problems. There are applications of TSP in operations research and engineering. A nice monograph [**51**] summaries various aspects of the work that has been done concerning TSP. See also expository articles [**49**], [**50**].

Finding the travelling salesman's shortest route to pass n cities in such a way that each city is visited exactly once represents the traditional formulation of TSP. It is assumed that non-negative distances d_{ij} between the cities i, j $(1 \leqslant i < j \leqslant n)$ are given and also that the travelling salesman starts his trip from an arbitrary city. If the travelling salesman does not return to the starting city, then the minimal traversed route is called an *open route* or simply a *path.*

This problem cannot be solved using derivatives. This is because the problem has a discrete character: we have to minimize a function defined on a finite set (the set of permutations of n cities in this case). Such problems belong to the area of combinatorial optimization. There is the obvious *brute force method* to solve such optimization problems: to calculate the value of the objective function for all points in the domain and to select minimum values. However, in the case of TSP and of many other combinatorial optimization problems the execution time of a brute force algorithm on best computers would last for thousands of years for quite modest dimensions of the problem instances (say a couple of dozens of cities in the case of TSP). Since applications require solving large scale problems, many "clever" algorithms and heuristics have been developed and a theory of complexity of algorithms and problems has been established.

One of most popular among algorithms which avoid total search is branch and bound. We shall describe branch and bound technique in a general framework with emphasis on the relevant details concerning the solving of the TSP.

For the sake of simplicity, we restrict ourselves to the following optimization (minimization or maximization) problem on weighted graphs (networks), which is still very general:

Let $\mathcal{A}$ be the set of all subgraphs of a graph G (with weights inherited from G). Let $\mathcal{F} \subseteq \mathcal{A}$ be the set of all subgraphs of G which posses some additional properties. The subgraphs from $\mathcal{F}$ are called *feasible.* We seek in $\mathcal{F}$ the elements with extremal (minimal or maximal) weights.

Let us assume that our optimization problem is a minimization problem. In the case of maximization the procedure would be similar.

In order to solve such a problem by a branch and bound algorithm, let $\mathcal{R}$ ($\mathcal{F} \subseteq \mathcal{R} \subseteq \mathcal{A}$) be a set of subgraphs for which there exists a polynomial time algorithm (say α) for finding the optimal element in $\mathcal{R}$. The set $\mathcal{R}$ corresponds to some relaxed variant of our problem (some feasibility conditions need not hold anymore).

To describe the algorithm (search procedure), we first introduce a *search tree* T as an auxiliary tool. T is a *rooted tree* with the root at a vertex r; all other vertices are the *descendants* of r. If f is any vertex, then its out-neighbors (called *sons* of f, f being their *father*) are denoted by $s_1, \ldots, s_n$. Each vertex, say f, corresponds to some subset $\mathcal{R}(f)$ of $\mathcal{R}$ and to a subproblem of the original problem (usually obtained by including and/or excluding some edges of G from the solution). The root r corresponds to the whole set $\mathcal{R}$. If f is a father and the solution of the relaxation task on the corresponding subproblem is not feasible and its length is smaller than the current lower bound (set at the beginning), then after branching at f by some branching rules (which "destroy" some "unfeasible details" in the solution of the relaxation task), the set $\mathcal{R}(f)$ is split into mutually disjoint subsets $\mathcal{R}(s_1), \ldots, \mathcal{R}(s_n)$ yealding new subproblems and new vertices in the search tree T. By solving the relaxation problem at some tree vertex with the use of the algorithm α, we obtain a lower bound for a feasible solution at this vertex. A global upper bound is provided at the beginning by taking any feasible subgraph (usually found by some quick heuristic). The branch and bound algorithm terminates when all subproblems in the search tree T are exhausted.

The above described general scheme of a branch and bound algorithm can be specified to solve the TSP by taking $\mathcal{F}$ to be the set of all Hamiltonian paths (or cycles or circuits – depending on the variant considered).

For a more detailed treatment of branch and bound algorithms see for example [**51**, pp. 361–401].

1.2. Graph Spectra and Other Graph Invariants. The *adjacency matrix* of a (multi)(di)graph G, with vertex set $\{1, 2, \ldots, n\}$, is the $n \times n$ matrix $A = (a_{ij})$ whose (i, j)-entry a_{ij} is equal to the number of edges, or arcs, originating at the vertex i and terminating at the vertex j. Two vertices of G are said to be *adjacent* if they are connected by an edge or arc. Unless we indicate otherwise we shall assume that G is an undirected graph without loops or multiple edges. The *degree* of a vertex is the number of vertices adjacent to that vertex.

The characteristic polynomial $\det(\lambda I - A)$ of the adjacency matrix A of G is called the *characteristic polynomial* of G and denoted by $P_G(\lambda)$. The eigenvalues of A (i.e., the zeros of $\det(\lambda I - A)$) and the spectrum of A (which consists of the n eigenvalues) are also called the *eigenvalues* and the *spectrum* of G, respectively. The spectrum of G is denoted by *spec* G. These notions are independent of vertex labelling because a reordering of vertices results in a similar adjacency matrix. The eigenvalues of G are usually denoted by $\lambda_1, \ldots, \lambda_n$; they are real because A is symmetric. Unless we indicate otherwise, we shall assume that $\lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_n$

and use the notation $\lambda_i = \lambda_i(G)$ for $i = 1, 2, \ldots, n$. Clearly, isomorphic graphs have the same spectrum.

The eigenvalues of A are the numbers λ satisfying $Ax = \lambda x$ for some non-zero vector $x \in \mathbb{R}^n$. Each such vector x is called an *eigenvector* of the matrix A (or of the labelled graph G) belonging to the eigenvalue λ. The relation $Ax = \lambda x$ can be interpreted in the following way: if $x = (x_1, x_2, \ldots, x_n)^T$, then $\lambda x_u = \sum_{v \sim u} x_v$ where the summation is over all neighbours v of the vertex u. If λ is an eigenvalue of A then the set $\{x \in \mathbb{R}^n : Ax = \lambda x\}$ is a subspace of $\mathbb{R}^n$, called the *eigenspace* of λ and denoted by $\mathcal{E}(\lambda)$ or $\mathcal{E}_A(\lambda)$. Such eigenspaces are called eigenspaces of G. Of course, relabelling of the vertices in G will result in a permutation of coordinates in eigenvectors (and eigenspaces).

The largest eigenvalue λ_1 of a graph G is called the emphindex of G; since adjacency matrices are non-negative there is a corresponding eigenvector whose entries are all non-negative.

Next we present certain notation, definitions and results from graph theory.

As usual, K_n, C_n and P_n denote respectively the *complete graph*, the *cycle* and the *path* on n vertices.

mG denotes the union of m disjoint copies of G. We write $V(G)$ for the vertex set of G, and $E(G)$ for the edge set of G.

If uv is an edge of G we write $G - uv$ for the graph obtained from G by deleting uv. For $v \in V(G)$, $G - v$ denotes the graph obtained from G by deleting the vertex v and all edges incident with v. More generally, for $U \subseteq V(G)$, $G - U$ is the subgraph of G induced by $V(G) \setminus U$.

A function defined on a family $\mathcal{G}$ of graphs is called a *graph invariant* for graphs in $\mathcal{G}$ if it is the same for isomorphic graphs in $\mathcal{G}$. Usually, graph invariants are numbers (integers, reals, etc.) but can be more complex objects (families of numbers, vectors, matrices, etc.).

Highly informative graph invariants from the title have not been defined precisely; we use this term informally. We shall say that a graph invariant is *highly informative* if it can be obtained quickly (possibly by a polynomial time algorithm) and if it contains a lot of information on the graph structure. It would be desirable that the invariant fully determines the graph (up to isomorphism as it is usually said). Such invariants would be obviously useful in solving the *graph isomorphism problem*, i.e., the problem of deciding whether or not two given graphs are isomorphic.

Let us consider some examples of graph invariants

1. *Vertex degrees.* The family of vertex degrees can be quickly calculated. However, the degree of a vertex is a kind of local invariant; it does not depend on the structure of the whole graph. Only neighbors of the vertex in question contribute to the value of its degree. It is not surprising that the family of vertex degrees does not say much on the graph structure, i.e. usually there are several graphs having a given family of vertex degrees. For example, a graph on 8 vertices having all vertex degrees equal to 2 can be one of the following three graphs: C_8, $C_5 \cup C_3$, $C_4 \cup C_4$.

2. *Spectrum.* Family of graph eigenvalues is obtained by considering extremal values of the Rayleigh quotient of the adjacency matrix. Eigenvalues depend in general case on all details on graph structure. Therefore more can be said on graph structure in the case that we know graph eigenvalues than in the case of knowing vertex degrees. Let us analyze the situation with graphs in which the vertex degrees are equal to 2. Such graphs are called *regular graphs of degree* 2.

Regular graphs of degree 2 are unions of cycles. One can verify by direct calculation that eigenvalues of the cycle C_n are real parts of the n-th roots of 2^n, i.e.,

$$spec\, C_n = \{Re\, \sqrt[n]{2^n}\} = \left\{2\cos\frac{2\pi}{n}j \;\middle|\; j = 0, 1, \ldots, n-1\right\}$$

The largest eigenvalue is $\lambda_1 = 2$ ($j = 0$) and the next one is two-fold: $\lambda_2 = \lambda_3 = 2\cos\frac{2\pi}{n}$ (for $j = 1$ and $j = n-1$). Suppose now that $G = \bigcup_{i=1}^{k} C_{n_i}$. Then

$$spec\, G = \bigcup_{i=1}^{k} \left\{2\cos\frac{2\pi}{n_i}j \;\middle|\; j = 0, 1, \ldots, n_i - 1\right\}$$

Given $spec\, G$, we can first establish that G is regular (by Theorem 3.22 of [**31**]) of degree 2. This is already information contained in the family of vertex degrees. But here we have more. Finding the second largest eigenvalue in modulus in $spec\, G$, we can determine the size n_i of the largest cycle in G. Gradually, by analyzing the whole spectrum we can determine the sizes of all cycles of G, i.e., determine G up to isomorphism.

In this way we have proved the following theorem (see [**12**] or [**31**, p. 167]).

THEOREM 1.1. *A regular graph of degree* 2 *is characterized by its spectrum.*

The reader might think that unions of cycles are not so interesting graphs to justify the space devoted to their spectral characterizations. However, the importance of this theorem will be shown in Subsection 2.1.

It seems that graph theoretical invariants, which contain a lot of information about the graph structure and thus are useful for the graph isomorphism problem, are obtained by solving some kind of optimization problem. Eigenvalues are also obtained in this way (as extrema of the Rayleigh quotient). The same holds for angles of a graph [**17**]. See 3.3 for other examples.

3. *A binary number.* A graph G can be characterized by the largest (or least) binary number obtained by concatenation of rows (or rows of the upper triangle) of adjacency matrices of G. The ordering of vertices which yields the characterizing binary number can be considered as a canonical vertex ordering. One can consider several variations of this idea but it turns out that the known algorithms for finding the graph characterizing quantity are exponential (cf. [**62**], [**4**]). Here a high price has been paid. We have an invariant which tells everything about the graph but it is time consuming to determine it. (However, this does not mean that under certain circumstances the extremal binary number has not been successfully used in recognizing graphs).

From the point of view of practical computation it is not very important to decide whether the graph isomorphism problem is NP-complete or belongs to P. Experience has shown that any reasonable algorithm for graph isomorphism testing performs well in average. However, the problem has great theoretical significance. Leaving aside the implications in the theory of complexity of algorithms and problems, one can say that the understanding of the kind of difficulties arising in the graph isomorphism problem enables the understanding of difficulties that appear in treating graph theory problems in general.

After having got equanted with these three examples we might be inclined to believe that spectral type graph invariants represent a good compromise between different reqirements on graph invariants. Therefore we describe another variant in defining graph eigenvalues.

Let $G = (V, E)$ be an undirected simple graph, where $V = \{1, \dots, n\}$ is the set of vertices and E is the set of edges. The *Laplacian* $L(G)$ of graph G is a symmetric matrix defined as $L(G) = D(G) - A(G)$, where $D(G)$ is the diagonal matrix with vertex degrees on the diagonal and $A(G)$ is the adjacency matrix of G.

The matrix $L(G)$ is positive semidefinite. If $\mu_1 \leqslant \dots \leqslant \mu_n$ are eigenvalues of $L(G)$, then $\mu_1 = 0$ with the corresponding eigenvector $e = (1, \dots, 1)$. All other eigenvalues have eigenvectors which belong to the set

$$S = \left\{ x = (x_1, \dots, x_n) \in \mathbb{R}^n \;\middle|\; \sum_{i=1}^{n} x_i = 0,\ \sum_{i=1}^{n} x_i^2 = 1 \right\}$$

According to Fiedler, the second smallest eigenvalue μ_2 of $L(G)$, is called the *algebraic connectivity* of G and denoted by $a(G)$. In [**37**] the following results are proved:

THEOREM 1.2. *The algebraic connectivity $a(G)$ has the properties:*

(*i*) $a(G) = \min\limits_{x \in S} x^T L(G) x$

(*ii*) $a(G) \geqslant 0$, $a(G) > 0$ *if and only if G is connected.*

Fiedler shows that the notion of the Laplacian and the algebraic connectivity can be generalized to graphs with positively weighted edges.

A C-edge-weighted graph $G_C = (V, E, C)$ is defined by a graph $G = (V, E)$ and a symmetric nonnegative weight matrix C such that $c_{ij} > 0$ if and only if $\{i, j\} \in E$. Now the Laplacian $L(G_C)$ is defined as $L(G_C) = \operatorname{diag}(r_1, \dots r_n) - C$, where r_i is the sum of the i-th row of C. The Laplacian $L(G_C)$ has similar characteristics as $L(G)$. Namely it is symmetric, positive semidefinite with the smallest eigenvalue $\mu_1 = 0$ and the corresponding eigenvector e. As before, the algebraic connectivity $a(G_C)$ is the second smallest eigenvalue of $L(G_C)$, which enjoys similar properties to those in Theorem 1.2.

THEOREM 1.3. (M. Fiedler [**37**]) *The generalized algebraic connectivity $a(G_C)$ has the following properties:*

(*i*) $a(G_C) = \min\limits_{x \in S} x^T L(G_C) x$

(*ii*) $a(G_C) \geqslant 0$, $a(G_C) > 0$ *if and only if G_C is connected.*

1.3. Semidefinite Programming. Semidefinite programming (SDP) has been one of the most active research areas in mathematical programming during the last decade. It is related to minimization of a linear function on the set of positive semidefinite matrices subject to linear constraints.

Recall that a symmetric matrix is called *positive semidefinite* (*positive definite*) if its eigenvalues are nonnegative (positive).

In order to define a semidefinite program, we need to introduce the appropriate notation. Let $S^{n\times n}$ denote the set of symmetric $n\times n$ matrices and let $S_+^{n\times n}$ denote the set of positive semidefinite $n\times n$ matrices, Then $S_+^{n\times n}$ is a closed convex cone in $\mathbb{R}^{n\times n}$ of dimension $n(n-1)/2$. We write $X \succcurlyeq 0$ ($X \succ 0$) to denote that X is a symmetric positive semidefinite (positive definite) matrix, and we write $X \succcurlyeq Y$ to denote that $X - Y \succcurlyeq 0$. For $A, B \in \mathbb{R}^{n\times n}$ the Frobeinus product is defined by

$$A \circ B = tr(A^T B) = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij} b_{ij}$$

If $A, B \in S^{n\times n}$ it follows that $A \circ B = tr(AB)$.

If $A, B \in S_+^{n\times n}$ it can be proved that $A \circ B \succcurlyeq 0$ and that $A \circ B = 0$ implies $AB = 0$ (see [**65**]).

Now a semidefinite program (SDP) can be formulated as:

$$\begin{array}{lll} & \textit{minimize} & C \circ X \\ (1) & \textit{subject to} & A_i \circ X = b_i,\ i = 1,\dots,m \\ & & X \succcurlyeq 0 \end{array}$$

where C, $A_1,\dots,A_m \in S^{n\times n}$, $b = (b_1,\dots,b_m) \in \mathbb{R}^m$ are given parameters and the unknown $n\times n$ matrix X is symmetric positive semidefinite. In the sequel P and P° will denote the feasible set of problem (1) and its relative interior, i.e.,

$$P = \{X \in \mathbb{R}^{n\times n} \mid A_i \circ X = b_i,\ i = 1,\dots,m,\ X \succcurlyeq 0\}$$
$$P^\circ = \{X \in \mathbb{R}^{n\times n} \mid A_i \circ X = b_i,\ i = 1,\dots,m,\ X \succ 0\}$$

Without loss of generality we may assume that matrices $A_1,\dots,A_m$ are linearly independent. It is easy to see that then (1) can be written in the form

$$\begin{array}{lll} & \textit{minimize} & c_0 + c^T z \\ (2) & \textit{subject to} & F_0 + \sum_{i=1}^{p} z_i F_i \succcurlyeq 0 \end{array}$$

where $z \in \mathbb{R}^p$ is the unknown vector, $p = n(n+1)/2 - m$, and $F_i \in S^{n\times n}$, $i = 0,\dots,p$, $c_0 \in R$, $c \in \mathbb{R}^p$ are the corresponding parameters. Indeed, problem (1) has n^2 scalar variables and $m + n(n-1)/2$ linear equations (m given explicitly and $n(n-1)/2$ following from the fact that X is symmetric). Hence there are $n^2 - m - n(n-1)/2 = n(n+1)/2 - m = p$ free variables which uniquely determine the remaining ones, i.e., there exist (symmetric) matrices F_0, $F_1,\dots,F_p$ such that

$$\{X \in S^{n\times n} \mid A_i \circ X = b_i, i = 1,\dots,m\} = \{X = F_0 + z_1 F_1 + \dots + z_p F_p \mid z \in \mathbb{R}^p\}$$

This implies that the feasible sets of (1) and (2) are equal. Moreover,

$$C \circ X = C \circ (F_0 + z_1 F_1 + \cdots + z_p F_p) = C \circ F_0 + z_1 C \circ F_1 + \cdots + z_p C \circ F_p$$

and we can take $c_0 = C \circ F_0$, $c = (C \circ F_1, \ldots, C \circ F_p)$.

Theoretical properties of the SDP problem have been studied in sixties, seventies and early eighties by several authors, e.g. Bellman, Fan [**7**], Craven, Mond [**11**], Fletcher [**38**], Rockafellar [**63**] Wolkowicz [**67**], etc. We shall state here only the main results. The dual problem associated to (1) is the following SDP problem of the type (2):

$$\begin{aligned} &\textit{maximize} \quad b^T y \\ &\textit{subject to} \quad \sum_{i=1}^{m} y_i A_i \leqslant C, \end{aligned}$$

which can be equivalently reformulated as:

$$\begin{aligned} &\textit{maximize} && b^T y \\ &\textit{subject to} && \sum_{i=1}^{m} y_i A_i + Z = C \\ &&& Z \geqslant 0 \end{aligned} \tag{3}$$

The feasible set of (3) and its relative interior will be denoted by D and D°, respectively, i.e.

$$D = \left\{ (Z,y) \in \mathbb{R}^{n\times n} \times \mathbb{R}^m \;\Big|\; \sum_{i=1}^{m} y_i A_i + Z = C,\ Z \geqslant 0 \right\}$$

$$D^\circ = \left\{ (Z,y) \in \mathbb{R}^{n\times n} \times \mathbb{R}^m \;\Big|\; \sum_{i=1}^{m} y_i A_i + Z = C,\ Z > 0 \right\}$$

It is easy to prove the following week duality theorem.

THEOREM 1.4. *If $X \in P$, $(Z,y) \in D$, then $C \circ X \geqslant b^T y$.*

PROOF. We have

$$C \circ X = \sum_{i=1}^{m} y_i A_i \circ X + Z \circ X = \sum_{i=1}^{m} y_i b_i + Z \circ X = b^T y + Z \circ X \tag{4}$$

As Z, $X \in S_+^{n\times n}$ it follows that $Z \circ X \geqslant 0$ and (4) implies $C \circ X \geqslant b^T y$. □

Let p^* and d^* be the optimal values of primal (1) and dual (3), i.e.,

$$p^* = \inf_{X \in P} C \circ X, \qquad d^* = \sup_{(Z,y)\in D} b^T y$$

Theorem 1.4 implies $p^* \geqslant d^*$. Let P^* and D^* be the corresponding sets of optimal solutions, i.e.,

$$P^* = \{X \in P \mid C \circ X = p^*\}, \qquad D^* = \{(Z,y) \mid b^T y = d^*\}.$$

It is easy to construct examples demonstrating that the sets P^* (D^*) can be empty even if p^* (d^*) is finite, which is not the case in linear programming. The next theorem gives conditions which guarantee that P^* and D^* are nonempty and that the duality gap $p^* - d^*$ is equal to zero.

THEOREM 1.5. *(i) Suppose that one of the following conditions hold:* $1°$ $P° \neq \emptyset$, $2°$ $D° \neq \emptyset$. *Then* $p^* = d^*$.

(ii) Suppose that $1°$ *and* $2°$ *hold. Then* $P^* \neq \emptyset$, $D^* \neq \emptyset$.

The proof is an application of the duality theory from convex analysis (see e.g., **[59]**, **[63]**).

If both conditions $1°$ and $2°$ hold it is easy to see that the set $P^* \times D^*$ is equal to the set of solutions to the system

$$(5) \qquad \begin{aligned} &\text{(a)} \quad X\,Z = 0 \\ &\text{(b)} \quad \sum_{i=1}^{m} y_i\, A_i + Z - C = 0 \\ &\text{(c)} \quad A_i \circ X - b_i = 0,\ i = 1,\dots,m \\ &\text{(d)} \quad X \succcurlyeq 0,\ Z \succcurlyeq 0 \end{aligned}$$

Indeed, if X and (Z, y) are optimal solutions of problems (1) and (3) their feasibility implies conditions (5b)–(5d). Moreover, $C \circ X = p^* = d^* = b^T y$. Since by (4) $Z \circ X = C \circ X - b^T y$ it follows that $Z \circ X = 0$, which implies $X\,Z = 0$, i.e., (5a) holds.

Let now $(X, Z, y) \in \mathbb{R}^{n\times n} \times \mathbb{R}^{n\times n} \times \mathbb{R}^m$ be a solution of (5a)–(5d). Then X and (Z, y) are feasible solutions of (1) and (3) and hence, by Theorem 1, $C \circ X \geqslant p^* \geqslant d^* \geqslant b^T y$. As $X\,Z = 0$ implies $Z \circ X = 0$ from (4) it follows $C \circ X = b^T y$, i.e., $C \circ X = p^*$, $b^T y = d^*$.

A strong impulse to further development of semidefinite programming was given by Nesterov and Nemirovski in a series of papers **[55, 56, 57, 58, 59]** written between 1988 and 1991 and by Alizadeh **[2]**, who have shown independently that interior point methods for linear programming can be directly extended to SDP. For example, the parametrized logarithmic barrier problem for linear programming extends to SDP as:

$$(6) \qquad \begin{aligned} &\textit{minimize} && C \circ X - t \ln(\det X) \\ &\textit{subject to} && A_i \circ X = b_i,\ i = 1,\dots,m \\ &&& X \succ 0 \end{aligned}$$

where $\ln(\det X)$ replaces the logarithmic barrier function $\sum_{j=1}^{n} \ln x_j$. The optimality conditions for this problem can be written as

$$(7) \qquad \begin{aligned} &\text{(a)} \quad X\,Z - tI = 0 \\ &\text{(b)} \quad \sum_{i=1}^{m} y_i\, A_i + Z - C = 0 \\ &\text{(c)} \quad A_i \circ X - b_i = 0,\ i = 1,\dots,m \\ &\text{(d)} \quad X \succ 0,\ Z \succ 0 \end{aligned}$$

which in fact is a parametrization of optimality conditions (5a)–(5d). Under the assumptions 1° and 2° from Theorem 1.5 it can be shown that for each $t > 0$ system (7a)–(7d) has the unique solution (X_t, Z_t, y_t). Moreover, $\lim_{t\to 0+}(X_t, Z_t, y_t) = (X^*, Z^*, y^*)$, where X^* solves (1) and (Z^*, y^*) solves (3) (for a proof see [**46**]).

The key idea of interior-point methods for SDP is to use Newton method in order to get approximate solutions of the parametrized system (7a)–(7d). A typical algorithm can be described as follows:

Algorithm:
Input: $X_0 \in P^\circ$, $(Z_0, y_0) \in D^\circ$, $\varepsilon > 0$
Initialization: Set $k = 0$, $t_0 = X_0 \circ Z_0/n$
Repeat until $t_k < \varepsilon$ do
(1) Set in (7a)–(7d) $t = t_k$
(2) Compute the Newton direction $(\Delta X_k, \Delta Z_k, \Delta y_k)$ at (X_k, Z_k, y_k).
(3) Choose $\alpha_k > 0$ such that
$(X_{k+1}, Z_{k+1}, y_{k+1}) = (X_k, Z_k, y_k) + \alpha_k(\Delta X_k, \Delta Z_k, \Delta y_k) \in P^\circ \times D^\circ$
(4) Set $t_{k+1} = X_{k+1} \circ Z_{k+1}/n$, $k \leftarrow k + 1$
End.

It should be noted that (7a) can be represented in many different ways, including for example $(X\,Z + Z\,X)/2 - t\,I = 0$, resulting in many different nonequivalent Newton directions, and hence different SDP methods. In terms of theoretical performance, the best SDP methods are guaranteed to reduce duality gap of the iterates by a fixed proportion in $O(\sqrt{n})$ iterations. This is identical to the complexity result for linear programming with n variables, even though the number of scalar variables in SDP is much larger (there are $n(n+1)/2$ entries in the symmetric matrix X). More precisely, the algorithm stops in $O(\sqrt{n}\log\frac{X_0 \circ Z_0}{n\epsilon})$ iterations, while the complexity of a single iteration of the algorithm is typically $O(\max\{m^2n^2, mn^3, m^3\})$. This gives the overall complexity bound $O(\max\{m^2n^{2.5}, mn^{3.5}, m^3n^{o.5}\})$.

There are many active research areas in semidefinite programming varying from development of different interior point algorithms and investigating their properties to writing efficient SDP codes capable of handling large sparse SDP problems. Special attention is payed to applications of SDP, which are very wide. The types of constraints that can be modelled in the SDP framework include linear inequalities, convex quadratic inequalities, lower bounds on matrix norms, lower bounds on determinants of symmetric positive semidefinite matrices, lower bounds on the geometric mean of a nonnegative vector, etc. Using these and other constructions the following problems can be stated as SDP problems: optimizing a convex quadratic form subject to convex quadratic inequalities, minimizing the volume of an ellipsoid that covers a given set of points and ellipsoids, maximizing the volume of an ellipsoid that is contained in a given polytope, a variety of maximum eigenvalue and minimum eigenvalue problems, etc. In particular, there is a growing interest in applications of SDP in combinatorial optimization where it is used in order to get satisfactory lower bounds on the optimal objective function value. Some examples are SDP relaxations for the max-cut problem, graph coloring problem and the travelling salesman problem. The next section gives a detailed description of SDP approach to the travelling salesman problem. A comprehensive survey of theory, algorithms and applications of semidefinite programming can be found in a recently published monograph [68].

2. Using Spectral Invariants in Problems of Combinatorial Optimization

In this section we elaborate Assertion 2 by describing the use of algebraic connectivity of a graph in solving TSP and in some clustering problems. In Subsection 2.1 we describe a model of TSP based on semidefinite programming and algebraic connectivity of graphs. Another way of using graph spectra in treating TSP is given in Subsection 2.2, where we introduce complexity indices for TSP. Subsection 2.3 describes some problems of clustering binary vectors and provides another example of using highly informative graph invariants in solving optimization problems.

2.1. Discrete Semidefinite Programming Model for TSP. Let $G = (V, E)$ be a complete undirected graph, where, as before, $V = \{1, \dots, n\}$ is the set of vertices and E is the set of edges. To each edge $\{i, j\} \in E$ a distance (cost) d_{ij} is associated such that the distance matrix $D = (d_{ij})_{n\times n}$ is symmetric and $d_{ii} = 0$, $i = 1, \dots, n$. Now the symmetric travelling salesman problem (TSP) can be formulated as follows: *find a Hamiltonian circuit of G with minimal cost.*

Algebraic connectivity of a Hamiltonian circuit is well known in the theory of graph spectra (see e.g. [31]). Since the graph is regular of degree 2, we have $L = 2I - A$. Hence, the Laplacian of a circuit with n vertices has the spectrum

$$2 - 2\cos(2\pi j/n),\ j = 1, \dots, n$$

and the second smallest eigenvalue is obtained for $j = 1$ and $j = n - 1$, i.e., $\mu_2 = \mu_3 = 2 - 2\cos(2\pi/n)$. This value will be denoted by h_n, i.e., $h_n = 2 - 2\cos(2\pi/n)$.

Now, Theorem 1.1 of Section 1.2 will be transformed into a form which is very useful in solving TSP. The next theorem, which gives a basis for the discrete semidefinite programming model of TSP, has been proved in [24] as a consequence of a more general result. For the sake of completeness we supply here a self-contained proof following [25].

THEOREM 2.1. *Let H be a spanning subgraph of G such that $d(i) = 2$, $i = 1, \dots, n$, where $d(i)$ is the degree of vertex i with respect to H, and let $L(H) = (l_{ij})_{n\times n}$ be the corresponding Laplacian. Let α and β be real parameters such that $\alpha > h_n/n$, $0 < \beta \leqslant h_n$. Then H is a Hamiltonian circuit if and only if the matrix $X = L(H) + \alpha J - \beta I$ is positive semidefinite, where J is the $n \times n$ matrix with all entries equal to one and I is the unit matrix of order n.*

PROOF. Let $0 = \mu_1 \leqslant \mu_2 \leqslant \cdots \leqslant \mu_n$ be the eigenvalues of $L(H)$ and let $x^1 = e$ and $x^i \in S, i = 2, \dots, n$, be the corresponding eigenvectors which form a basis for $\mathbb{R}^n$. It is easy to check that J has two eigenvalues: 0, with multiplicity $n-1$ and the corresponding eigenvectors $x^2, \dots, x^n$, and n with e as its eigenvector. Therefore

$$Xe = (L + \alpha J - \beta I)e = (\alpha n - \beta)e$$

$$Xx^i = (L + \alpha J - \beta I)x^i = (\mu_i - \beta)x^i,\ i = 2, \dots, n$$

which means that $\alpha n - \beta$ and $\mu_i - \beta$, $i = 2, \ldots, n$ are eigenvalues of X with eigenvectors $e, x^2, \ldots, x^n$, respectively.

The conditions of Theorem 2.1 garantee that H is a 2-matching, i.e., it is either a Hamiltonian circuit or a collection of at least two disjoint subcircuits. In the first case $\mu_2 = h_n$, while in the second, according to Theorem 1.2, $\mu_2 = 0$. As $\alpha > h_n/n$ in both cases it follows that $\alpha n - \beta > \mu_2 - \beta$, i.e., the smallest eigenvalue of X is equal to $\mu_2 - \beta$.

Suppose that H is a Hamiltonian circuit. Then $\beta \leqslant h_n$ implies $\mu_2 - \beta = h_n - \beta \geqslant 0$, i.e., matrix X is positive semidefinite. Suppose now that X is positive semidefinite. Then $\mu_2 - \beta \geqslant 0$ and $\beta > 0$ imply $\mu_2 = a(H) > 0$ and by Theorem 1.2 it follows that H is a connected 2-matching, i.e., a Hamiltonian circuit. □

It follows from Theorem 2.1 that a spanning subgraph H of G is a Hamiltonian circuit if and only if its Laplacian $L(H)$ satisfies the following conditions:

$$l_{ii} = 2, \ i = 1, \ldots, n \tag{8}$$

$$X = L(H) + \alpha J - \beta I \text{ is positive semidefinite}, \ \alpha > h_n/n, \ 0 < \beta \leqslant h_n \tag{9}$$

Starting from (8) and (9) the following discrete semidefinite programming model of TSP can be defined

$$\textit{minimize } F(X) = \sum_{i=1}^{n} \sum_{j=1}^{n} \Big(-\frac{1}{2} d_{ij} \Big) x_{ij} + \frac{\alpha}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} d_{ij} \tag{10}$$

subject to

$$x_{ii} = 2 + \alpha - \beta, \ i = 1, \ldots, n \tag{11}$$

$$\sum_{j=1}^{n} x_{ij} = n\alpha - \beta, \ i = 1, \ldots, n \tag{12}$$

$$x_{ij} \in \{\alpha - 1, \alpha\}, \ i, j = 1, \ldots, n, i < j \tag{13}$$

$$X \geqslant 0, \tag{14}$$

where $X \geqslant 0$ denotes that the matrix $X = (x_{ij})_{n \times n}$ is symmetric and positive semidefinite and α and β are chosen according to Theorem 2.1. Matrix $L = X + \beta I - \alpha J$ represents the Laplacian of a Hamiltonian circuit if and only if X satisfies (11)–(14). Indeed, constraints (11)–(13) provide that L has the form of a Laplacian with diagonal entries equal to 2, while condition (14) guarantees that L corresponds to a Hamiltonian circuit. Therefore, if X^* is an optimal solution of problem (10)–(14), then $L^* = X^* + \beta I - \alpha J$ is the Laplacian of an optimal Hamiltonian circuit of G with the objective function value $\sum_{i=1}^{n} \sum_{j=1}^{n} (-\frac{1}{2} d_{ij}) l_{ij}^* = F(X^*)$.

The well-known integer programming formulation of TSP reads:

$$\textit{minimize } \sum_{i \in V} \sum_{j > i} d_{ij} x_{ij} \tag{15}$$

subject to

$$\sum_{j<i} x_{ji} + \sum_{j>i} x_{ij} = 2, \qquad i \in V\,; \tag{16}$$

$$\sum_{i\in S}\sum_{\substack{j\in S\\ j>i}} x_{ij} \leqslant |S| - 1, \quad \text{for all } S \subset V,\ S \neq \emptyset; \tag{17}$$

$$x_{ij} = 0 \text{ or } 1 \qquad i,j \in V,\ j > i. \tag{18}$$

The subtour elimination inequalities (17) can also be written as

$$\sum_{i\in S}\sum_{\substack{j\in V-S\\ j>i}} x_{ij} + \sum_{i\in V-S}\sum_{\substack{j\in S\\ j>i}} x_{ij} \geqslant 2, \qquad \text{for all } S \subset V,\ S \neq \emptyset.$$

Each of n constraints in group (16) requires exactly two edges to be incident to every vertex and constraints of type (17) are subtour elimination constraints excluding the subtours with less than n vertices.

It is important to note that in our discrete semidefinite programming model of TSP (10)–(14), the single condition (14) replaces all subtour elimination constraints in the standard integer programming model!

A natural semidefinite relaxation of the travelling salesman problem is obtained when discrete conditions (13) are replaced by inequality conditions:

$$\textit{minimize } F(X) \tag{19}$$

subject to

$$x_{ii} = 2 + \alpha - \beta,\ i = 1, \dots, n \tag{20}$$

$$\sum_{j=1}^{n} x_{ij} = n\alpha - \beta,\ i = 1, \dots, n \tag{21}$$

$$\alpha - 1 \leqslant x_{ij} \leqslant \alpha,\ i,j = 1, \dots, n,\ i < j \tag{22}$$

$$X \succcurlyeq 0 \tag{23}$$

It is easy to see that the relaxation (19)–(23) can be expressed in the standard form of an SDP problem. Indeed, constraint (20) can be represented as $A_i \circ X = 2 + \alpha - \beta$, where $\circ$ is the Frobenius product and A_i is a symmetric $n \times n$ matrix with 1 at the position (i,i) and all other entries equal to 0. Similarly, condition (21) is equivalent to $B_i \circ X = 2(n\alpha - \beta)$, where B_i has 2 at the position (i,i) while all the remaining elements of the i-th row and the i-th column are equal to 1, and all the other entries are zero. Finally, condition (22) can be expressed as $2(\alpha - 1) \leqslant C_{ij} \circ X \leqslant 2\alpha$, where C_{ij} has 1 at the positions (i,j) and (j,i) and zero otherwise. Since SDP problem (19)–(23) depends on parameters α and β it represents a class of semidefinite relaxations of TSP. In the sequel, members of this class will be referred to as SDP relaxations.

Let us denote by P and P° the feasible set of problem (19)–(23) and its relative interior. For each $X \in P$ the corresponding Laplacian $L = X + \beta I - \alpha J$ can be interpreted as the Laplacian of the weighted graph $G_L = (V, E_L, C_L)$, where

$E_L = \{\{i,j\} \in E \mid l_{ij} < 0\}$ and $C_L = 2I - L$. If α and β satisfy the conditions of Theorem 2.1 then, using similar arguments as in the proof of Theorem 2.1, it can be shown that $X \geqslant 0$ is equivalent to $a(G_L) \geqslant \beta$ (see also [24]). Hence, by Theorem 1.3 graph G_L is connected. It immediately follows that 2-matchings with disjoint subcircuits cannot correspond to any X in P.

It is easy to see that $P^\circ \neq \emptyset$. Indeed, if e.g. $\hat{L} = \left(2 + \frac{2}{n-1}\right)I - \frac{2}{n-1}J$, then $\hat{X} = \hat{L} + \alpha J - \beta I = \left(2 + \frac{2}{n-1} - \beta\right)I + \left(\alpha - \frac{2}{n-1}\right)J$ has the eigenvalues $2 + \frac{2}{n-1} - \beta$ with the multiplicity $n-1$ and $n\alpha - \beta$ with the multiplicity 1. Since $n\alpha - \beta > 0$ and $2 + \frac{2}{n-1} - \beta \geqslant 2 + \frac{2}{n-1} - h_n > 0$ for $n \geqslant 4$, it follows that $\hat{X} \in P^\circ$, $n \geqslant 4$.

For $\beta < h_n$ matrices X which correspond to Laplacians of Hamiltonian circuits are in P°, while for $\beta = h_n$ these matrices belong to $P \setminus P^\circ$. It is clear that the best relaxation is obtained for $\beta = h_n$. Concerning the parameter α, it is always sufficient to choose $\alpha = 1$.

The semidefinite relaxation (19)–(23) is substantially different from the existing TSP relaxations. It should be pointed out that it cannot be theoretically compared neither with 2-matching nor with 1-tree. Indeed, if we consider TSP model (10)–(14) it is easy to see that X which corresponds to the Laplacian of a 2-matching satisfies (11)–(13) but need not satisfy (14). In the case of 1-tree, the condition (11) is relaxed, while (12), (13) and (14) hold (see [24]). Preliminary numerical experiments on randomly generated problems with $10 \leqslant n \leqslant 20$ which are reported in [24], indicate that SDP relaxation gives considerably better lower bounds than both 1-tree and 2-matching.

We have implemented two branch and bound algorithms with the SDP relaxation (with $\alpha = 1$, $\beta = h_n$) and one with the 1-tree relaxation. The last one was implemented to check the correctness of the results. All algorithms are based on the general branch and bound scheme as described in [51]. We used a FORTRAN implementation of the branch and bound shell from the package TSP-SOLVER [21], [29]. An initial upper bound was obtained in all cases by the 3-optimal heuristic. The depth first search was used to select the next subproblem.

The two branch and bound algorithms differ only in their branching rules (the first one defined by Vollgnant and Jonker, see [25]):

Algorithm 1. At the first vertex of degree greater than 2 in the weighted graph representing the SDP solution an edge is excluded in each son;

Algorithm 2. The first non-integer entry of the SDP solution matrix is replaced in the sons by 0 and 1 respectively.

Inequality conditions (22) were handled adding $n^2 - n$ slack variables each represented by a 1×1 block as accepted by the software.

For solving the SDP relaxation tasks we used a modification of CSDP 2.3 software package developed by Borchers [8, 9] in C language. The package is based on a predictor–corrector variant of the interior point algorithm presented by

Helmberg, Rendl, Vanderbei, Wolkowicz [**43**]. The experiments were performed on an Alpha 800 5/400 computer. Preliminary computational results were reported in [**25**]. A part of the results is presented in the next subsection. Further numerical evidence with larger TSP instances is given in [**47**], [**48**].

2.2. Complexity Indices for TSP. In this section we will illustrate how some invariants related to the solution of SDP relaxation (19)–(23) could be implemented as complexity indices in an adaptive solution approach for TSP. Such an approach, introduced in [**29**], is based on the following principles:

For a given branch and bound (B&B) algorithm and a given maximal number of relaxation tasks R_S which are allowed to be solved within the algorithm, TSP instances are classified into two classes: *hard* and *easy* instances. The hard class contains TSP instances for which solving more than R_S relaxation tasks is required to reach an optimal solution, while the easy class consists of the remaining instances.

Since hard instances usually require a lot of computing time, there is some interest to recognize such instances before the algorithm starts. If a concrete instance is recognized to be hard, instead of finding an exact solution, a suboptimal solution could be found by an efficient heuristic.

The recognition of hard and easy instances is realized using the notion of complexity indices.

Given an instance of the TSP, *the instance complexity* of this instance for the given B&B algorithm can be defined as *the number of relaxation tasks which need to be solved within the applied algorithm to reach an optimal solution.*

Any number assigned to an instance which contains some information on the instance complexity (with respect to a given B&B algorithm) will be called a *complexity index.*

Usually, complexity index is a numerical graph invariant of a (weighted) graph related to the solution of the relaxation task for the instance considered. In the context of this study, special attention will be paid to highly informative graph invariants, since we might expect that just these will serve as good complexity indices.

Here we assume that there exists an efficient (polynomial) algorithm for determining the index under the consideration.

Since an instance complexity of the TSP for a given B&B algorithm is related to the number of relaxation tasks, it is reasonable to determine the value of a complexity index on the basis of solved relaxation tasks within the algorithm. This is based on the expectation that the branch and bound algorithm will run for longer, the more relaxation solution(s) are distanced from an optimal solution, and that this information could be extracted from one or several relaxation solution(s). Each type of a possible relaxation used in some variant of B&B algorithm offers a variety of complexity indices. In this way complexity indices depend upon a B&B algorithm and so special complexity indices for each variant of a B&B algorithm can be introduced.

There are no theoretical results described in the literature which would indicate the existence of efficient complexity indices for a particular instance of NP-hard problems, in spite of the fact that this would be of obvious practical importance. As a mater of fact, we do not see how a theory of complexity indices for instances of NP-hard problems could be set up based on known results.

The idea of complexity indices has been initiated in [**54**] in relation to the TSP. The indices offered have been intuitively justified and their validity supported by some experimental results. The largest eigenvalue of the adjacency matrix of a minimal spanning tree has been introduced in [**15**]] as a complexity index for the travelling salesman problem and its validity supported by some results from the theory of graph spectra [**31**].

As a selection criterion for the most informative index, the measure of statistical dependence of the value of the complexity index and the number of the solved relaxation tasks is used. This measure ought to reflect as much as possible the extent and the type of the dependence. It is also pointed out in [**15**] that the efficiency of complexity indices is related to statistical distribution of the set of instances which are intended to be solved.

Let $\mathcal{N}$ be a set of the TSP instances defined by distance matrices with elements (i.e., arc lengths) from a given distribution. Let the output of the applied branch and bound algorithm be presented by two sequences of real numbers (b_i) and (c_i), $i = 1, 2, \ldots, |\mathcal{N}|$, where b_i is the number of solved relaxation tasks and c_i is the value of the corresponding complexity index, both referring to the i-th instance of the TSP in the set $\mathcal{N}$. Under this assumptions we can interpret sequences (b_i) and (c_i) as the realizations of random variables B and C in a statistical experiment.

The measure of dependence of a complexity index and the number of solved relaxation tasks can be interpreted as a degree of dependence of the random variables C and B and estimated by the methods of correlation analysis.

The coefficient of linear correlation for two sequences (b_i) and (c_i) is defined by

$$C_{BC} = \frac{1}{\vartheta_B \vartheta_C} \sum_{i=1}^{|\mathcal{N}|} (b_i - \bar{m}_B)(c_i - \bar{m}_C),$$

where $\bar{m}_B$, $\bar{m}_C$ and ϑ_B, ϑ_C are mean values and variances of the corresponding sequences (b_i) and (c_i), respectively.

Under the assumption that the random variables B and C obey the normal distribution, the correlation coefficient C_{BC} is a reliable estimation of linear dependence of the random variables B and C.

The efficiency of the complexity index can be statistically estimated measuring the linear correlation between the index value and the number of relaxation tasks solved within the B&B algorithm.

Several invariants can be considered as complexity indices for the TSP with respect to B&B algorithms based on SDP relaxation [**26**]:

Let X be the solution of SDP relaxation (19)–(23) and $L = X + h_n I - J$. Then L determines the weighted graph $W_L = (V, E_L, C_L)$, where $E_L = \{\{i, j\} \in E \mid l_{ij} < 0\}$ and $C_L = 2I - L$, the corresponding unweighted graph $G_L = (V, E_L)$ and a stochastic matrix $S_L = I - \frac{1}{2}L$. The most efficient indices introduced in [**26**] are the following:

I_1: the number of edges of G_L

I_2: the second smallest eigenvalue of the Laplacian of G_L

I_3: the entropy of S_L, i.e., value equal to $\sum\limits_{(i,j)\in E_L} \frac{l_{ij}}{2} \log_2 \left(- \frac{l_{ij}}{2} \right) - \frac{n}{2}$

I_4: $\sum\limits_{i=1}^{n} |\mu_i - \mu_i^*|$, where $\mu_1, \mu_2, \ldots, \mu_n$ and $\mu_1^*, \mu_2^*, \ldots, \mu_n^*$ are sequences of nondecreasing eigenvalues of the Laplacians of G_L and a Hamiltonian circuit, respectively.

I_5: the same sum as in I_4 but with eigenvalues of the Laplacian of W_L instead of G_L.

I_6: the number of vertices of the G_L with degrees greater than 2.

The efficiency of indices I_k, $k = 1, \ldots, 6$, has been investigated in [**26**]:

For each dimension 20, 25, 30, 35 we consider 50 randomly generated TSP instances with distances uniformly distributed in the interval [1,999]. To each instance one of B&B algorithms based on SDP relaxation is applied (see Subsection 2.1).

The coefficients of the linear correlation between values of indices $I_k (k = 1, \ldots, 6)$ and the number of relaxation tasks for dimensions n = 20, 25, 30, 35 are summarized in Table 1. Results indicate that the most reliable indices are I_1, I_4 and I_6 with almost significant correlation.

TABLE 1. Values of the linear correlation coefficients

index n	I_1	I_2	I_3	I_4	I_5	I_6
20	0.53	0.35	0.51	0.53	0.53	0.53
25	0.48	0.49	0.21	0.48	0.48	0.49
30	0.29	0.21	0.32	0.29	0.42	0.33
35	0.56	0.52	0.37	0.56	0.38	0.55
average value	0.47	0.39	0.35	0.47	0.45	0.48

Index I_4 is a spectrally based invariant and, having in view facts from Subsection 1.2, one would expect that it performs better than I_1 and I_6. The obtained experimental results presented in this subsection indicate the lack of theoretical explanations of phenomena with complexity indices, the need for experiments with instances of higher dimensions and, perhaps, the need for better classification of graph invariants than the intuitive approach, adopted in this study on highly informative graph invariants.

An idea how to improve the results with complexity indices is already given in another context in [**29**, pp. 23–25]. One can consider linear combinations of already defined complexity indices as well as the invariants of some short edge subgraphs of the input weighted graph.

We shall describe now our adaptive procedure for solving TSP.

The most important parameter in the adaptive solution approach for TSP is I^* which represents the estimated value of the complexity index I corresponding to the maximal allowed number of solved relaxation tasks R_S. In general R_S depends on the problem dimension n and this function is chosen by the user, while I^* depends on both n and statistical distribution of TSP instances. The value of I^* is used to classify instances in hard and easy classes in the following way [**29**]:

(1) easy instances, $I \leqslant I^*$, and (2) hard instances, $I > I^*$.

More precisely, when $I \leqslant I^*$, then the number of generated subproblems within the branch and bound procedure is expected to be less than R_S, while for $I > I^*$ this number is expected to be greater than R_S.

In the case when the solution process for TSP is based on SDP relaxation easy instances can be solved by one of B&B algorithms from Subsection 2.1, while hard instances are handled by a heuristic developed in [**27**]. The heuristic uses limited branching based on the number of edges with weights equal to one in the graph W_L. Namely, already mentioned experiments with a set of 50 randomly generated TSP instances for each of dimensions n=20, 25, 30, 35 show that high percentage of edges from W_L with weights equal to one stay in the optimal Hamiltonian circuit, which is illustrated in Table 2. This suggests the following modification of the branch and bound algorithms from Subsection 2.1. Starting from an initial upper bound obtained by the 3-opt heuristic and the solution X of the initial SDP relaxation, all edges from W_L with weights equal to one are fixed, and the branching is performed on the remaining edges.

TABLE 2

1	2	3	4
20	79.8%	97.7%	84%
25	84.1%	98.2%	78%
30	82.2%	98.1%	74%
35	85.4%	97.4%	60%

The proposed heuristic solves the TSP by limited branching and therefore it has in the worst case exponential complexity. Nevertheless it performs very well in practice, which is illustrated in Table 3. The test examples are the same as in Table 2.

The columns in Table 2 contain the following data:

(1) dimension n of TSP;
(2) the average percentage of the edges with weights equal to one in W_L w.r.t. n;
(3) the average percentage of edges from W_L with weights equal to one which stay in the optimal solution;
(4) the percentage of TSP instances for which all edges from W_L with weights equal to one stay in the optimal solution.

TABLE 3. Performance of the heuristic

1	2	3	4	5	6
20	93.6%	53.2%	0.14%	5	31
25	85.1%	59.6%	0.12%	7	80
30	79.2%	72.9%	0.37%	8	47
35	62.5%	89.6%	0.62%	10	92

The columns in Table 3 contain the following data:

(1) dimension of TSP;
(2) percentage of instances for which optimal solution was reached;
(3) percentage of instances for which the heuristic improved the initial upper bound obtained by 3-opt heuristic;
(4) average relative distance from the optimal objective function value $(((f_H - f_{\text{opt}})/f_{\text{opt}})\%)$;
(5) average number of subproblems solved within the heuristic;
(6) maximal number of subproblems solved within the heuristic.

In [**27**] the described adaptive solution procedure was tested on the same set of randomly generated instances used to generate Tables 1,2 and 3.

Before solving the instances we need first to decide upon the value of R_S-the maximal number of relaxation tasks which are allowed to be solved within the adaptive procedure. Here for each dimension we take that R_S is equal to the maximal number of subproblems solved within the heuristic given in column 6 of Table 3. As suggested in [**27**] the most reliable estimation for I_k^* was achieved by averaging those values in the correlation field which were equal (or approximatively equal) to R_S.

The performance of the adaptive solution procedure can be measured by two parameters f and t determined on the basis of the whole set of considered TSP instances. The parameter f represents the average percentage difference between the objective function value f_{ad} obtained by the adaptive procedure and the optimal objective function value f_{opt} determined by the B&B algorithm, i.e.

$$f = \frac{f_{\text{ad}} - f_{\text{opt}}}{f_{\text{opt}}}\%.$$

Value t is the "saving" of CPU-time, i.e., the percentage difference between the average number of solved relaxation tasks within the B&B and the adaptive procedure.

In order to measure the quality of the adaptive solution with one parameter we use the value $k = t/(1 + f/100)$ proposed in [**29**].

The adaptive procedure was tested on the same sets of TSP instances for each of complexity indices I_k, $k = 1, \ldots, 6$ as before. According to [**27**], the best performance was obtained for I_6.

In Table 4 we report on relevant details for this case.

TABLE 4. Results of the adaptive procedure for I_6

n	20	25	30	35
I_6^*	14	12	11	12
No. of hard instances	10	11	25	23
No. of detected hard instances	11	15	27	23
No. of easy instances	37	36	23	25
No. of detected easy instances	36	32	21	25
No. of incorrectly classified instances	7	16	10	10
exactness of decision	85.1%	66.00%	79.2%	79.2%
f	0.01%	0.05%	0.21%	0.21%
average No. of subproblems in adaptive algorithm	13	36	56	61
average No. of subproblems in optimal algorithm	23	69	198	238
t	43.5%	47.8%	71.7%	74.4%
k	43.5%	47.8%	71.6%	74.2%

TABLE 5

index	I_1	I_2	I_3	I_4	I_5	I_6
average exactness of decision	76.8%	73.7%	69.5%	76.3%	71.0%	77.4%
average value k	57.1%	46.4%	56.6%	56.1%	59.3%	59.3%

The results for the remaining complexity indices were also reasonable. Table 5 summarizes the most important indicators of the efficiency of the adaptive procedure: the average exactness of decision per dimension and the average value k per dimension.

On the basis of results presented in Tables 4 and 5 we can conclude that both the exactness of decision and the quality measure k are satisfactory for all complexity indices.

Another use of complexity indices in the TSP solving procedures is described in [**32**]. A new search strategy in B&B algorithms has been developed. The traditional backtrack strategies (depth-first search and breath-first search) are not optimal and therefore the so called *jumptrack* strategies have been considered in the literature. In such strategies any active subproblem can be selected following certain criteria. Usually, one selects a subproblem with smallest lower bound. Complexity indices have been introduced in [**32**] to help to select next subproblem. An ordered list of *interesting* subproblems has been introduced. The strategy takes the first subproblem from the list and branches it using the depth-first search until a complexity index starts to increase. Several variants of such a search strategy have been described in [**32**]. Computational results show better behaviour of these strategies compared with other ones.

2.3. Data Clustering. In this subsection we consider the problem of clustering data (see, e.g., [**1**], [**3**]). Clusters in some cases are obtained by solving some optimization problems. Again the algebraic connectivity of a graph is useful.

The algebraic connectivity is known to be a very useful parameter for describing the "shape" of a graph (see, e.g., [**31**, p. 266]). Indeed, low algebraic connectivity shows small connectivity and girth and high diameter, although such a statement lacks a precise formulation. In the context of clustering, low algebraic connectivity indicates that the graph has good clustering properties.

The data are usually represented by vectors from $\mathbb{R}^n$. Euclidean or other kind of distance function $d(x, y)$ is assumed to be defined for any $x, y \in \mathbb{R}^n$. Given a set of vectors from $\mathbb{R}^n$, the problem is to partition it into subsets called *clusters* under various conditions. Clustering methods are supposed to produce clusters which have the property that vectors from the same cluster in some sense are "closer" one to the other than the vectors from different clusters. The number of clusters may but need not to be given in advance. Sometimes cardinalities of clusters are given or limited by additional conditions.

There are difficulties in applying traditional clustering procedures to discrete data. We describe a graph theoretical approach in clustering binary vectors. New clustering procedures are combined from several algorithms and heuristics from graph theory and combinatorial optimization.

We consider clustering of discrete data. A typical example of discrete data are binary vectors, i.e. elements of B^n where $B = \{0, 1\}$. When standard clustering procedures (see, e.g., [**1**], [**3**]) are applied to binary vectors, the resulting clustering has usually a low quality. Among other things, the clustering is highly dependent of the ordering of vectors.

To avoid these difficulties it seems reasonable to use specific properties of discrete data and to apply combinatorial, including graph theoretical, tools in handling the problem. We have developed a number of complex graph theoretical procedures for clustering binary vectors [**16**], [**18**], [**20**]. See also [**22**] and [**61**].

A *hypercube* H_n of dimension n is the graph whose vertex set is B^n and two n-tuples are adjacent if they differ in exactly one coordinate. The number of coordinates in which n-tuples $x, y \in B^n$ differ is called the *Hamming distance* between x and y.

For a graph G we define its *k-th power* G^k. The graph G^k has the same vertex set as G and vertices x and y are adjacent in G^k if they are at (graph theoretical) distance at most k in G. For $k = 0$ the graph G^k consists of isolated vertices. For $k = 1$ we have $G^k = G$. If X is a subset of the vertex set of a graph G, then $G(X)$ denotes the subgraph of G induced by X.

Let $X \subset B^n$ be a set of binary vectors (n-tuples) which is to be clustered. Our procedures for clustering make use of the graph sequence

$$H_n^0(X),\ H_n^1(X),\ H_n^2(X), \ldots, H_n^n(X)$$

which is called the *basic graph sequence.*

Note that two vectors $x, y \in X$ are at the Hamming distance k if they are not adjacent in $H_n^{k-1}(X)$ and are adjacent in $H_n^k(X)$. For $i = 1, \dots, n$ the graph $H_n^i(X)$ has all edges from $H_n^{i-1}(X)$ plus those ones connecting vectors at *Hamming* distance i. $H_n^0(X)$ has only isolated vertices while $H_n^n(X)$ is a complete graph.

Let the vertex set X of a graph G be partitioned into subsets $X_1, X_2, \dots, X_m$. A *condensation* of G is a weighted graph on vertices $x_1, x_2, \dots, x_m$ (called *supervertices*) in which x_i and x_j are connected by an edge if there is at least one edge between X_i and X_j in G. Both supervertices and edges in the condensation carry weights. The weight of the supervertex x_i is equal to $|X_i|$ while the weight of the edge between x_i and x_j is equal to the number of edges between X_i and X_j. We consider a condensation as a multigraph where edge weights are interpreted as edge multiplicities while supervertices as vertices and supervertex weights are ignored.

In the clustering procedure, which will be described, some algorithms and heuristics, described in the literature, will be used. We shall define them here (see [**18**] for details).

Algorithm CP. This is an algorithms for finding components of a graph.

Algorithm JM. This is an algorithm for partitioning a connected (multi-) graph into two parts.

Heuristic KL. This is a heuristic for partitioning the vertex set of a (multi-) graph into two parts of given cardinalities with a minimum number of edges between vertices from different parts [**45**].

Let X be a set of binary vectors of dimension n and suppose we have to cluster it into k ($k > 1$) clusters. For $k = 2$ we consider the problem in two variants: 1° Cluster cardinalities are not given, 2° Cluster cardinalities are given.

Our procedure consists of two phases.

Phase 1. We form the basic graph sequence. Let c_i be the number of components of $H_n^i(X)$. Components are sequentially determined in graphs from the basic sequence by algorithm CP. We have $c_0 = |X| \geqslant c_1 \geqslant c_2 \geqslant \cdots \geqslant c_n = 1$.

There is a non-negative integer s such that $c_s \geqslant k > c_{s+1}$. If $c_s = k$, the components of $H_n^s(X)$ are clusters and the procedure is finished. If $c_s > k > c_{s+1}$ we proceed to Phase 2.

Phase 2. We distinguish cases: 1) $k = 2$ and 2) $k > 2$.

Case $k = 2$. Now $H_n^{s+1}(X)$ is connected and we consider the condensation of the graph $H_n^{s+1}(X)$ in which components of $H_n^s(X)$ play role of supervertices.

We consider two subcases:

1° Cluster cardinalities are not given;

2° Cluster cardinalities are given.

Subcase 1°. If $c_s > 10$ any of the following two procedures can be applied to the condensation of $H_n^{s+1}(X)$:

a) algorithm JM;

b) heuristic KL.

In any of these cases we get two clusters and the whole procedure is finished.

In variant b) the user can select the range of cluster cardinalities and the number of randomly generated starting clusterings. The result in variant a) can serve as a hint for the range of cluster cardinalities in variant b).

If $c_s \leqslant 10$, we form all partitions of the vertex set of the condensation of $H_n^{s+1}(X)$ into two parts since there are only $2^{c_s} - 2$ such partitions. We find the best partition with respect to a selected quality criterion (e.g., minimizing the edge number between two parts). The whole procedure is thus finished.

Subcase 2°. We apply algorithm JM to the condensation of $H_n^{s+1}(X)$. If the partition thus obtained shows cluster cardinalities required, we have done. Otherwise we apply heuristic KL to the graph H_n^{s+1} where the starting partitions are formed on the basis of information obtained by the working of the algorithm JM. Let p, q $(p \geqslant q)$ be the required cluster cardinalities. Let algorithm JM have given a solution with cluster cardinalities r, s $(r \geqslant s)$. If $p < r$, from the cluster of cardinality r we choose those p vertices for which moduli of the coordinates of the eigenvector from algorithm JM are as great as possible. If $p > r$, then $q < s$, and from the cluster of cardinality s we choose q vertices as above. The result of the working of heuristic KL for the starting partition so formed is compared with result for other, randomly generated, starting partitions.

Case $k > 2$. Now we have $c_s > k > c_{s+1}$ and we get a clustering into k clusters in one of the following two ways

1) by splitting some of c_{s+1} components of the graph $H_n^{s+1}(X)$ into parts;
2) by uniting some of c_s components of $H_n^s(X)$.

We use first way if k is closer to c_{s+1} than to c_s and the second one otherwise.

Splitting components we perform by partitioning a component into two parts and by iterating this procedure. First we partition components of $H_n^{s+1}(X)$ which do not exist in $H_n^s(X)$ and if there are not sufficiently many such components we treat sequentially those which exist in $H_n^i(X)$ and not in $H_n^{i-1}(X)$ for $i = s, s-1, \ldots$. For components of $H_n^i(X)$ $(i = s+1, s, \ldots)$ we form condensations with supervertices corresponding to components of $H_n^{i-1}(X)$ and for each condensation we determine the ratio of the algebraic connectivity and the number of vertices. Condensations are ordered by this ratio and partitioned sequentially into two parts starting from those with a smallest ratio. In each step of partitioning the newly generated components are treated as above. For partitioning components into two parts we apply the procedure from the case $k = 2$ above.

When uniting components we consider all possibilities of uniting if $c_s - k < 4$. Otherwise we apply the Ward method, which is one of the best hierarchical clustering methods (see, e.g., [**1**], [**3**]).

Algorithm JM and the calculation of the algebraic connectivity have complexity $O(|X|^3)$ while other parts of the procedure have lower complexities. Therefore the whole procedure has complexity $O(|X|^3)$ and this is the same as in many standard clustering procedures. However, theoretical reasons and numerical experiments show that the graph theoretical procedure is superior to standard procedures in clustering binary vectors.

3. Defining Graph Invariants by Solving Optimization Problems on Graphs

Section 3 elaborates Assertion 1. Subsection 3.1 introduces star bases and a canonical star basis of a graph. This basis is obtained by solving several optimization tasks and is useful in treating the graph isomorphism problem. It is shown in 3.2 that using canonical star basis one can construct a polynomial time algorithm for checking the isomorphism of graphs with bounded multiplicities of eigenvalues. Other examples of defining highly informative graph invariants are given in 3.3.

3.1. Star Partitions and Canonical Star Bases. Spectral techniques in graph theory are based on the eigenvalues of the adjacency and other graph matrices. These techniques have been further developed by considering, in addition, some invariants of eigenspaces of graphs, namely graph angles. Introduction of star partitions and canonical bases can be considered as a result of efforts in the same direction – to enrich spectral techniques in graph theory.

Let G be a graph with vertices $1, \ldots, n$ and $(0,1)$-adjacency matrix A. Let $\mu_1, \ldots, \mu_m$ ($\mu_1 > \cdots > \mu_m$) be the distinct eigenvalues of A, with corresponding eigenspaces $S_1, \ldots, S_m$. For each $i \in \{1, \ldots, m\}$, let k_i be the multiplicity of μ_i.

Let us consider the spectral decomposition of A:

$$A = \mu_1 P_1 + \cdots + \mu_m P_m.$$

Thus P_i represents the orthogonal projection onto S_i and, if $\{e_1, \ldots, e_n\}$ is the standard orthonormal basis of $\mathbb{R}^n$, the vectors $P_i e_1, \ldots, P_i e_n$ constitute a *eutactic star* (see [**33**]). Norms of vectors from these eutactic stars are *angles* of a graph. Rows (or columns) of P_i are vectors of the eutactic star associated with S_i. The Gram matrix of these vectors is just the matrix P_i.

A partition $X_1 \dot\cup \cdots \dot\cup X_m$ of the vertex set $\{1, \ldots, n\}$ is called a *star partition*, with *star cells* $X_1, \ldots, X_m$, if for each $i \in \{1, \ldots, m\}$ the vectors $P_i e_j$ ($j \in X_i$) are linearly independent. In this situation a comparison of dimensions shows that $|X_i| = k_i$ ($i = 1, \ldots, m$) and the vectors $P_i e_j$ ($j \in X_i$) form a basis B_i of S_i. Then $B_1 \cup \cdots \cup B_m$ is a basis of $\mathbb{R}^n$, called in [**33**] a *star basis* corresponding to A (a construction applicable to any symmetric matrix with real entries).

It was shown in [**33**] that the following three theorems hold.

THEOREM 3.1. *Every graph G has a star partition.*

THEOREM 3.2. *The partition $X_1 \dot\cup \cdots \dot\cup X_m$ is a star partition if and only if for each $i \in \{1, \ldots, n\}$, μ_i is not an eigenvalue of $G - X_i$.*

THEOREM 3.3. (Reconstruction theorem) *A graph G is uniquely determined by an eigenvalue μ_i, the subgraph $G - X_i$, and the subgraph $G - E(X_i)$ where X_i is a star cell belonging to μ_i and $E(X)$ is the set of edges of G whose both end points are in the set X.*

It has been shown in [**34**] that it is possible to construct one star partition of a graph G in time bounded by a polynomial function of the number n of vertices of the graph G. One approach is related to the Hungarian algorithm for constructing a perfect matching in a bipartite graph. The complexity of this approach is at most $O(n^4)$. The other approach uses matroid theory and a rough complexity estimation is $O(n^5)$.

Let S be the matrix of a basis $\mathcal{S}$ of eigenvectors (in particular, a star basis or even some kind of the canonical basis) of $\mathbb{R}^n$ corresponding to a graph G. Columns of S are vectors of $\mathcal{S}$. We have $S = (S_1 \cdots S_m)$, where S_i is the matrix of the star basis $\mathcal{S}_i$ of the eigenspace $\mathcal{E}(\mu_i)$. The adjacency matrix A can be expressed in the form $A = S\Lambda S^{-1}$ where Λ is a diagonal matrix with eigenvalues of A at the main diagonal. Putting $(S^{-1})^T = (X_1 \cdots X_m)$ and using relation $S^{-1}S = I$ we get $X_i = (S_i^T S_i)^{-1} S_i^T$. This is sufficient to prove the following proposition.

Proposition 1. We have $A = \mu_1 S_1 (S_1^T S_1)^{-1} S_1^T + \cdots + \mu_m S_m (S_m^T S_m)^{-1} S_m^T$.

Proposition 1 gives the spectral decomposition of the adjacency matrix A of a graph. Hence, $P_i = S_i(S_i^T S_i) S_i^T$ is the projection onto the i-th eigenspace $\mathcal{E}(\mu_i)$. The matrix $W_i = S_i^T S_i$ is the Gram matrix of the star basis $\mathcal{S}_i$ of the eigenspace $\mathcal{E}(\mu_i)$. (If the basis is orthonormal W_i is equal to a unit matrix).

If $\mathcal{S}$ is a precisely defined canonical basis, eigenvalues $\mu_1, \ldots, \mu_m$ and matrices $S_1, \ldots, S_m$ comprise the canonical code of a graph. All graph properties can be derived from the canonical code of the graph at least by reconstructing the adjacency matrix A and applying some graph theoretical algorithms. Of course, it would be of interest to derive procedures for a more direct link between the canonical code of the graph and the graph properties we are interested in.

The problem is how to select a canonical basis of eigenvectors. Without additional restrictions there are infinitely many such bases and if we want to select one by an optimization task, the set of feasible solutions is infinite and compact so that we are led to a problem of continual (global) optimization. However, if we restrict ourselves to star bases, we encounter a problem of finding an extremal value of a function defined on a finite set since the number of star bases is finite. Hence, we have to solve a problem of *combinatorial optimization*.

Let us note that the number of star bases of a graph, although finite, is very large since we have to consider star bases corresponding to all orders (permutations) of the vertex set.

Given a graph G on n vertices, the notion of a canonical basis of eigenvectors of G for $\mathbb{R}^n$ which is related to the notions of the *star basis* and the *star partition* of G has been introduced in [**33**]. This canonical basis is called the *canonical star basis.* The canonical star basis is unique for a graph i.e. it does not depend on the vertex labelling. Finding this canonical basis involves several extremal tasks similarly as in finding an extremal binary number (see Subsection 1.2). To construct this basis we have introduced a total order of graphs, called CGO (canonical graph ordering) and a quasi-order of vertices called CVO (canonical vertex ordering). Both CGO and CVO are defined recursively in terms of graphs with fewer than n vertices.

The definition of the canonical star basis enables the formulation of the following theorem (whose proof is obvious):

THEOREM 3.4. *Two graphs are isomorphic if and only if they have the same eigenvalues and the same canonical bases.*

Several improvements to the procedure from [**33**] have been proposed in [**19**]. See also [**35**]. One can show that the procedure is reduced in some parts to special cases of some well-known combinatorial optimization problems, such as the maximal matching problem, the minimal cut problem, the maximal clique problem, etc. This sheds some light on the algorithmical complexity of the procedure for finding a canonical basis, i.e., tells something about the graph isomorphism problem.

A procedure for testing the isomorphism of graphs, which is based on the spectral decomposition of matrices has been described in [**69**].

Since the canonical star basis together with eigenvalues of G determines G up to isomorphism, algorithms for finding the canonical basis and some of its variations are studied in [**19**]. The emphasis is given on the following three special cases: graphs with distinct eigenvalues, graphs with bounded eigenvalue multiplicities and strongly regular graphs. This technique provides another proof of a result of L. Babai et al. [**5**] that isomorphism testing for graphs with bounded eigenvalue multiplicities can be performed in polynomial time (see next subsection). One can show that the canonical basis in strongly regular graphs is related to the graph decomposition into two strongly regular induced subgraphs (these decompositions are described in [**41**]). Examples of distinguishing between cospectral strongly regular graphs by means of the canonical basis are provided. The behaviour of star partitions of regular graphs under operations of complementation and switching is studied in [**19**] as well.

The canonical star basis (and star bases in general) can be very useful in studying other problems.

3.2. The Maximal Clique Problem and Bounded Multiplicities. The procedure of finding the canonical star basis can be designed so that it contains a kind of the maximal clique problem. This is especially useful in the case of graphs with bounded multiplicities of eigenvalues. This subsection is mainly written following [**19**].

A star basis of a graph G with distinct eigenvalues $\mu_1, \ldots, , \mu_m$ of multiplicities $k_1, \ldots, k_m$ is characterized by weighted graphs $W_1, \ldots, W_m$ of orders $k_1, \ldots, k_m$, respectively (see the comment after Proposition 1 in the previous subsection). In *orthodox* star bases (i.e., bases among which the canonical star basis is selected) the sequence $W_1, \ldots, W_m$ is lexicographically maximal using ordering of weighted graphs specially defined in [**33**], [**19**], [**35**], i.e., canonical weighted graphs ordering (CWGO) and canonical weighted graphs vertex ordering (CWGVO). Instead of finding several (or all) star bases and selecting maximal ones among them, we can find maximal sequences $W_1, \ldots, W_m$ and check whether star bases with such sequences exist.

Let us assume that G is a connected graph. Then μ_1 is a simple eigenvalue and W_1 is reduced to squares of coordinates of the eigenvector belonging to μ_1. As proved in [**34**] any vertex can form a star cell X_1 for μ_1. Hence, for W_1 we select the square of a coordinate of the eigenvector belonging to μ_1 with a maximal modulus.

Next we have to select X_2, the star cell belonging to μ_2. That means we have to find a principal submatrix W_2 of P_2 of order k_2 so that the weighted graph determined by W_2 is maximal. Now we can consider P_2 as a weighted graph in which we have to find a clique of order k_2 with a maximal weight if we define the weight of a clique just to be the matrix W_2. We decide which of two given cliques has greater weight in accordance with ordering of weighted graphs (CWGO). The complexity of this decision depends on the order of the clique.

Finding a maximal clique in a weighted graph has been considered in [6]. The problem is essentially similar to the problem of finding a (maximal) clique in a graph without weights on edges [42]. Roughly speaking, we have to check all $\binom{n}{k_i}$ principal submatrices of order k_i. If k_i is fixed, $\binom{n}{k_i}$ is a polynomial in n of degree k_i, i.e., $\binom{n}{k_i} = O(n^{k_i})$. If the size of the clique is not restricted, the problem of finding a maximal clique (decision version) is known to be NP-complete. Note that $\binom{n}{cn}$ for a fixed c $(0 < c < 1)$ is not polynomially bounded.

Once we have found X_2 such that W_2 is maximal we can decide easily whether a star partition exists in which X_2 is a cell. (A necessary but not sufficient condition for this is that the graph $G - X_2$ does not have an eigenvalue μ_2. An example is provided by the Petersen graph; there is no star partition in which $X_1 \cup X_2$ induces C_6.)

More generally, given any partially built partition we can in a polynomial time, using algorithms from [34], extend it to a star partition or establish that this cannot be done.

It should be noted that our reductions of the graph isomorphism problem to some well known combinatorial optimization problems do not involve general cases of these problems; in fact, we have special cases determined by special features of weight matrices in question (eigenvector and projector matrices). This is important especially in the case when the general problem is NP-complete (NP-hard) as in the case of the problem of finding a maximal clique. Note that such reductions of the graph isomorphism problem to special cases of NP-hard problems (of unknown complexities) have been already noticed elsewhere (see [64] where a special case of the maximal clique problem occurs).

It has been proved by L. Babai et al. [5] that isomorphism testing for graphs with bounded multiplicities of eigenvalues can be performed in a polynomial time.

Using above ideas we can confirm this result.

If eigenvalue multiplicities are bounded by an absolute constant a, then the size of the maximal clique we have to find is limited also to at most a. It is known that a maximal clique of limited size can be found in a polynomial time, i.e., in time $O(n^a)$ in this case. We can in a polynomial time examine and keep information on all induced subgraphs whose vertex sets have cardinalities equal to eigenvalue multiplicities.

Hence we can find an orthodox star basis in a polynomial time. In fact, we can find all orthodox star bases in a polynomial time and go on in finding quasi-canonical bases and the canonical basis.

Ordering vertices in star cells of orthodox star bases by corresponding CWGVOs can be done in a time bounded by a function of the constant a. We can imagine that for testing isomorphism of graphs in which each eigenvalue has multiplicity at most a we have prepared a (finite) table of automorphism groups and CVO for graphs with at most a vertices. (It is even possible to use all orderings-at most $a!$-of vectors in an orthodox star basis.) Hence, all maximization procedures, needed to find the canonical star basis, can be performed in polynomial time and the result by L. Babai et al. follows.

Remark 1. The result can be extended to graphs (as noted also in [**5**]) in which all but one eigenvalue have bounded multiplicities and this property is hereditary (holds also for any induced subraph). The hereditarity of the property in question was not assumed in [**5**]. Perhaps it can be avoided also here but we have assumed it since the multiplicity of an eigenvalue can increase when going from graphs to subgraphs and the subgraph induced by the star cell corresponding to the eigenvalue of unbounded multiplicity can have more than one eigenvalue above the bound for eigenvalues. In this case we modify the notion of an orthodox star basis in such a way that the matrix W_i corresponding to the eigenvalue μ_i, whose multiplicity is not bounded, is put at the end of the original matrix sequence $W_1, \ldots, W_m$ which should be lexicographically maximal (i.e., we have now the sequence $W_1, \ldots, W_{i-1}, W_{i+1}, \ldots, W_n, W_i$). Namely, we readily find in polynomial time star cells corresponding to matrices $W_1, \ldots, W_{i-1}, W_{i+1}, \ldots, W_n$ while the cell of unbounded size, corresponding to W_i is determined by the vertices which remain. It is also not necessary to order vertices in this star cell; it is sufficient to use the Reconstruction theorem (Theorem 3.3 from Subsection 3.1).

Remark 2. Let us finally note that the graph isomorphism problem can be also reduced to a problem of finding a certain kind of matching in an auxiliary bipartite graph. However, the number of vertices of this bipartite graph depends on multiplicities of eigenvalues. The graph in question is the incidence graph between the set of distinct eigenvalues of the original graph G and the set of subsets of the vertex set of G with cardinalities equal to multiplicities of eigenvalues. There is an edge between vertices representing an eigenvalue μ and a subset X of vertices if and only if $|X|$ is equal to the multiplicity of μ and $G - X$ does not have an eigenvalue μ. A star partition of G is represented by a matching which satisfies some additional requirements but we shall not go into details.

As indicated, using canonical star bases we can relate the graph isomorphism problem to some problems of combinatorial optimization (the problem of finding a maximal matching, the maximal clique problem, the problem of finding a bipartition with an extremal number of edges between the parts, etc.). Some of these problems can be solved in polynomial time while the others are known to be NP-hard. Arguments pro and contra the existence of a polynomial algorithm for the graph isomorphism problem both exist.

3.3. Other Highly Informative Graph Invariants. In this final subsection we shortly report on some other graph invariants which can be considered as highly informative.

Spectra of weighted adjacency matrices have served to introduce a new important graph invariant in [**66**]. For a connected graph G we introduce the class $\mathring{A}_G$ of matrices $A = (a_{ij})$ for which $a_{ij} > 0$ if i and j are adjacent and $a_{ij} = 0$ otherwise. Let $\mu_1, \mu_2, \ldots, \mu_m$ ($\mu_1 > \mu_2 > \cdots > \mu_m$) be distinct eigenvalues of A with multiplicities $k_1 = 1, k_2, \ldots, k_m$, respectively. Let $\mu(G) = \max k_2$, where maximum is taken over the class $\mathring{A}_G$. For example, $\mu(K_n) = n-1$ and $\mu(K_{3,3}) = 4$. It is proved that G is planar if and only if $\mu(G) \leqslant 3$. It is conjectured that $\mu(G) \geqslant \chi(G) - 1$, where $\chi(G)$ is the chromatic number of G. The validity of this conjecture would imply the four color theorem!

The Lovász *theta function* has been introduced in [**53**] when solving a long standing problem in information theory. The theta function can be defined in many equivalent ways: via an extremal problem concerning eigenvalues of graphs, via a semidefinite programming model and in some other ways. For a short review on this important graph invariant see [**39**].

See [**10**] for other interesting graph invariants.

References

[1] D. Acketa, *Selected Topics in Pattern Recognition with Applications*, (in Serbian), Mathematical Institute, Novi Sad, 1986

[2] F. Alizadeh, *Combinational optimization with interior point methods and semidefinite matrices*, Ph.D. Thesis, University of Minnesota, 1991

[3] M. R. Anderberg, *Cluster Analysis for Applications*, Academic Press, New York, 1973

[4] L. Babai, *On the complexity of canonical labelling of strongly regular graphs*, SIAM J. Computing 9 (1980), 212–216

[5] L. Babai, D. Yu. Grigorjev, D. M. Mount, *Isomorphism of graphs with bounded eigenvalue multiplicity*, Proc. 14th Annual ACM Symposium on Theory of Computing, San Francisko 1982, New York, 1982, 310–324

[6] E. Balas, V. Chvatal, J. Nešetril, *On the maximum weight clique problem*, Mathematics of Operation Research, 12:3 (1987), 522–535

[7] R. Bellman, K. Fan, *On systems of linear inequalities in Hermitean matrix variables*, V. Klee, ed., Proc. Symposium in Pure Mathematics, Amer. Math. Soc., Providence, RI, 1963, 1–11

[8] B. Borchers, *CSDP, a C library for semidefinite programming*, Optimization Methods and Software 11-12 (1999), 613–623

[9] B. Borchers, *CSDP 2.3 user's guide*, Optimization Methods and Software 11-12 (1999), 597–611

[10] N. S. Chaudhari, D. B. Phatak, *A new approach for graph isomorphism*, J. Combin. Inform. System Sci. 13:1-2 (1988), 1–19

[11] B. Craven, B. Mond, *Linear programming with matrix variables*, Linear Algebra Appl. 38 (1981), 73–80

[12] D. Cvetković, *Graphs and their spectra*. Univ. Beograd, Publ. Elektrotehn. Fak. Ser. Mat. Fiz., No. 354 - No. 356 (1971), 1–50

[13] D. Cvetković, *The main part of the spectrum, divisors and switching of graphs*, Publ. Inst. Math. (Beograd) 23(37) (1978), 31–38

[14] D. Cvetković, *Combinatorial Matrix Theory with Applications to Electrical Engineering, Chemistry and Physics*, (Serbo-Croatian), Naučna knjiga, Beograd, 1980

[15] D. Cvetković, *Adaptive solving of the travelling salesman problem*, (Serbian), unpublished report, Faculty of Electrical Engineering, Belgrade, 1987, 1–48

[16] D. Cvetković, *Combinatorial algorithms and heuristics for clustering points of a hypercube, I, II, III, IV.* (in Serbian), University of Belgrade, Faculty of Electrical Engineering, Belgrade, pp. 32 + 39 + 53 + 34, unpublished report, 1988

[17] D. Cvetković, *Some graph invariants based on the eigenvectors of the adjacency matrix*, R. Tošić, D. Acketa, V. Petrović, R. Doroslovački, eds., Graph Theory, Proc. Eighth Yugoslav Seminar on Graph Theory, Novi Sad, April 17-18, 1987, Univ. Novi Sad, Inst. Math., Novi Sad, 1989, 31–42

[18] D. Cvetković, *Graph theoretical procedures in clustering discrete data*, Univ. Beograd, Publ. Elektrotehn. Fak., Ser. Mat. 3 (1992), 21–26

[19] D. Cvetković, *Star partitions and the graph isomorphism problem*, Linear and Multilinear Algebra 39 (1995), 109–132

[20] D. Cvetković, *A graph theoretical procedure for clustering binary vectors*, Ars Combinatoria 46(1997), 267–276

[21] D. Cvetković, M. Čangalović, V. Dimitrijević, L. Kraus, M. Milosavljević, S. Simić, *TSP-SOLVER-A programming package for the traveling salesman problem*. Univ. Beograd, Publ. Elektrotehn. Fak. Ser. Mat. 1 (1990), 41–47

[22] D. Cvetković, M. Čangalović, Dj. Dugošija, V. Kovačević-Vujčić, S. Simić, J. Vuleta, red. D. Cvetković, V. Kovačević-Vujčić, *Combinatorial Optimization*, (in Serbian), *Matematička teorija i algoritmi*, Društvo operacionih istraživača Jugoslavije, Beograd, 1996.

[23] D. Cvetković, M.Čangalović, V. Kovačević-Vujčić, *Semidefinite programming and traveling salesman problem*, R. Petrović, D. Radojević, eds., Proceedings of XXV Yugoslav Symposium on Operations Research, Herceg Novi, Yugoslavia, 1998, 239–242

[24] D. Cvetković, M.Čangalović, V. Kovačević-Vujčić, *Semidefinite relaxations of the traveling salesman problem*, YUJOR 9:2 (1999), 157–168

[25] D. Cvetković, M. Čangalović, V. Kovačević-Vujčić, *Semidefinite programming methods for the symmetric traveling salesman problem*, G. Cornuejols, R. E. Burkard, G. J. Woeginger, eds., Integer Programming and Combinatorial Optimization, Proc. 7th Internat. IPCO Conf., Graz, Austria, June 1999, Lecture Notes Comp. Sci. 1610, Springer-Verlag, Berlin, 1999, 126–136

[26] D. Cvetković, M.Čangalović, V. Kovačević-Vujčić, *Complexity indices for the traveling salesman problem based on a semidefinite relaxation*, M. Vujošević, M. Martić, eds., Proceedings of XXVI Yugoslav Symposium on Operations Research, Beograd, Yugoslavia, 1999, 177–180

[27] D. Cvetković, M. Čangalović, V. Kovačević-Vujčić, *Adaptive approach to the traveling salesman problem using a semidefinite relaxation*, S. Cvetić, S. Guberinić, eds., Proceedings of XXVII Yugoslav Symposium on Operations Research, Beograd, Yugoslavia, 2000, 213–216

[28] D. Cvetković, V. Dimitrijević, M. Milosavljević, *A spectral traveling salesman problem complexity index*, (in Serbian), Zbornik radova XI bosansko-hercegovačkog simpozijuma iz informatike, Sarajevo-Jahorina '87, 1987, knj. 2, 276-1–276-5

[29] D. Cvetković, V. Dimitrijević, M. Milosavljević, *Variations on the Travelling Salesman Theme*, Libra Produkt, Belgrade, 1996

[30] D. Cvetković, M. Doob, I. Gutman, A. Torgašev, *Recent Results in the Theory of Graph Spectra*, North-Holland, Amsterdam, 1988

[31] D. Cvetković, M. Doob, H. Sachs, *Spectra of Graphs*, 3rd ed., Johann Ambrosius Barth, Heidelberg-Leipzig, 1995

[32] D. Cvetković, S. Mitrović-Minić, *A search strategy in branch and bound algorithms for the traveling salesman problem*, E. Krčmar-Nožić, ed., Proceedings of XX Yugoslav Symposium on Operations Research, Beograd, 1993, 137–140

[33] D. Cvetković, P. Rowlinson, S. Simić, *A study of eigenspaces of graphs*, Linear Algebra Appl. 182 (1993), 45–66

[34] D. Cvetković, P. Rowlinson, S. Simić, *On some algorithmic investigations of star partitions of graphs*, Discrete Appl. Math. 62 (1995),119–130

[35] D. Cvetković, P. Rowlinson, S. Simić, *Eigenspaces of Graphs*, Cambridge University Press, Cambridge, 1997

[36] M. Fiedler, *Algebraic connectivity of graphs*, Czechoslovak Math. J. 23 (1973), 298–305

[37] M. Fiedler, *Laplacian of graphs and algebraic connectivity*, Combinatorics and Graph Theory, vol. 25, Banach Center Publications, PWN-Polish Scientific Publishers, Warsaw, 1989, 57–70

[38] R. Fletcher, *Semidefinite matrix constraints in optimization*, SIAM J. Control Optim. 23 (1985), 493–513

[39] M. Goemans, *Semidefinite programming in combinatorial optimization*, Math. Program. 79 (1997), 143–161

[40] M. X. Goemans, D. P. Williamson, *Improved approximation algorithms for maximum cut and satisfability problems using semidefinite programming*, J. ACM 42 (1995), 1115–1145

[41] W. H. Haemers, D. G. Higman, *Strongly regular graphs with strongly regular decomposition*, Linear Algebra Appl. 114/115 (1989), 379–398

[42] P. Hansen, N. Mladenović, *A comparison of algorithms for the maximum clique problem*, YUJOR 2:1 (1992), 3–11

[43] C. Helmberg, F. Rendl, R. J. Vanderbei, H. Wolkowicz, *An interior point method for semidefinite programming*, SIAM J. Optimization 6 (1996), 342–361

[44] D. Karger, R. Motwani, M. Sudan, *Approximate graph coloring by semidefinite programming*, J. ACM 45 (1998), 246–265

[45] B. W. Kernighan, S. Lin, *An efficcient procedure for partioning graphs*, Bell Sys. Tech. J. 49 (1970), 291–307

[46] M. Kojima, S. Shindoh, S. Hara, *Interior point methods for the monotone semidefinite linear complementarity problem in symmetric matrices*, SIAM J. Optimization 7 (1997), 86–125

[47] V. Kovačević-Vujčić, M. Čangalović, J. Kratica, *Solving a semidefinite relaxation of the traveling salesman problem*, Central European J. Oper. Res. 10:4 (2002), 277–296.

[48] V. Kovačević-Vujčić, M. Čangalović, J. Kratica, *A predicter-correcter algorithm for a semidefinite relaxation of the traveling salesman problem*, J. Todorović, M. Vujošević, B. Vasić, eds., Proceedings of XXIX Yugoslav Symposium on Operations Research, Tara, 2002, IX13–IX16.

[49] G. Laporte, *The traveling salesman problem: An overview of exact and approximate algorithms*, European J. Operational Research 59 (1992), 231–247

[50] G. Laporte, *Exact algorithms for the traveling salesman problem and the vehicle routing problem*, Les Cahiers du GERAD G-98-37, July, 1998

[51] E. L. Lawler, J. K. Lenstra, A. H. G. Rinnooy Kan, D. B. Shmoys, *The Traveling Salesman Problem*, John Wiley& Sons, Chichester-New York-Brisbane-Toronto-Singapore, 1985

[52] J. H. van Lint, R. M. Wilson, *A Course in Combinatorics*, Cambridge University Press, Cambridge, 1992

[53] L. Lovász, *On the Shannon capacity of a graph*, IEEE Trans. Inform. Theory 25 (1979), 1–7

[54] M. Milosavljević, V. Dimitrijević, *A complexity analysis concerning the exact solution of the travelling salesman problem*, (in Serbian), Zbornik radova X bosansko-hercegovačkog simpozijuma iz informatike, Sarajevo-Jahorina '86, 1986, knj. 2, 271-1–271-6

[55] Y. Nesterov, A. Nemirovski, *Polynomial time barrier methods in convex programming*, Ekonomika i Matem. Metody 24 (1988), 1084–1091

[56] Y. Nesterov, A. Nemirovski, *Self-concordant functions and polynomial time methods in convex programming*, USSR Acad. Sci., Central Economic&Mathematical Institute, Moscow, 1989

[57] Y. Nesterov, A. Nemirovski, *Optimization over positive semidefinite matrices*, USSR Acad. Sci., Central Economic&Mathematical Institute, Moscow, 1990

[58] Y. Nesterov, A. Nemirovski, *Conic formulation of a convex programming problem and duality*, USSR Acad. Sci., Central Economic&Mathematical Institute, Moscow, 1991

[59] Y. Nesterov, A. Nemirovski, *Interior-Point Polynomial Methods in Convex Programming*, Studies in Apllied Mathematics, vol. 13, SIAM, Philadelphia, 1994

[60] C. H. Papadimitrou, K. Steiglitz, *Combinatorial Optimization; Algorithms and Complexity*, Prentice-Hall, Englewood Cliffs, New Jersey, 1982

[61] S. Petrović, *Combinatorial algorithms and heuristics for clustering vertices of a graph and applications to pattern recognition*, (in Serbian), Master Thesis, University of Belgrade, Faculty of Electrical Engineering, Beograd, 1988

[62] R. C. Read, D. R. Corneil, *The graph isomorphism disease*, J. Graph Theory 1 (1977), 339–363

[63] R. T. Rockafellar, *Convex Analysis*, Princeton University Press, Princeton, 1970

[64] V. G. Timkovski, *Some recognition problems that are connected with the isomorphism of graphs*, (in Russian), Kibernetika (Kiev), 1988, no. 2, 13-18, 133; translated in Cybernetics 24:2 (1988), 153–160

[65] L. Vandenberghe, S. Boyd, *Semidefinite programming*, SIAM Review 38 (1996), 49–95

[66] Y. C. de Verdiére, *Sur un novel invariant des graphes et un critére planarité*, J. Combinatorial Theory B 50 (1990), 11–21

[67] H. Wolkowicz, *Some applications of optimization in matrix theory*, Linear Algebra Appl. 40 (1981), 101–118

[68] H. Wolkowicz, R. Saigal, L. Vandenberghe, *Handbook of Semidefinite Programming; Theory, Algorithms, and Applications*, Kluwer Academic Publishers, Boston-Dordrecht-London, 2000

[69] T. Živković, *On graph isomorphism and graph automorphism*, J. Math. Chem. 8:1-3 (1991), 19–37

Miodrag Mateljević

DIRICHLET'S PRINCIPLE, UNIQUENESS OF HARMONIC MAPS AND EXTREMAL QC MAPPINGS

CONTENTS

Preface

This expository paper consists of the various uniqueness theorems which follow, in general, from the length-area principle of Grötzsch. The structure of this paper is as follows. In Section I we give the main ideas and basic results. In the subsections A, B, C, D and E we discuss connections between the Grötzsch principle, Teichmüller's approach, the Main Inequality and Dirichlet's principle. In the subsections F and G we consider extremal problems for quasiconformal mappings. In particular, we give short review of new results and solve some problems, which originally were subject of investigation of Teichmüller, Reich, Strebel and the other mathematicians.

In Section II we give the outline of proofs of some properties of harmonic maps, using different tools: Dirichlet's principle, minimizing sequences, different versions of Reich–Strebel inequality. We also give a proof of well-known Beurling theorem.

In Section III using Lemma C1 we prove the inequality of Reich and Strebel for Riemann surfaces of finite analytic type and new version of an inequality of Reich and Strebel. We use this result to study the uniqueness properties of harmonic mappings (see section II).

Section IV is an extended version of the lecture given by the author at VIII Romanian–Finish Seminar, Iassy, August '99.

Recently, in [MM1], [BMM] and [BLMM], characterizations of unique extremality and example of unique extremal dilatation of nonconstant modulus have been obtained.

In this section our primary purpose is to give a short exposition of some of the main result of the authors' joint papers, mentioned above, and sketch further progress in the study of a more general concept. We also discuss the Beltrami equation, and we show the necessity of the Hamilton–Krushkal condition.

The most part of the paper consists of the lectures communicated by the author and the other members of the Seminar, at the University of Belgrade during several last years. The author also talked extensively about this subject in a number of places.

I. Introduction

A. The problem of Grötzsch. If Q is a square and R is a rectangle, not a square, there is no conformal mapping of Q on R which maps vertices on vertices. Instead, Grötzsch asked for the most nearly conformal mapping of this kind and took the first step toward the creation of a theory of quasiconformal mappings.

Let $w = f(z)$ be a mapping from one region to another. Recall that we use notation $df = p\,dz + q\,d\bar{z}$, where $p = \partial f$ and $q = \bar{\partial} f$. The complex (Beltrami) dilatation is $\mu_f = \mathrm{Belt}\,[f] = q/p$. The dilatation of f is

$$D_f = \frac{|p| + |q|}{|p| - |q|}.$$

We pass to the Grötzsch problem and give it a precise meaning by saying that f is most nearly conformal if $\sup D_f$ is as small as possible.

Let R, R' be two rectangles with sides a, b and a', b'. We may assume that $K = \frac{b'}{a'} : \frac{b}{a} \geqslant 1$. The mapping f is supposed to be C^1-homeomorphism from $\overline{R}$ onto $\overline{R'}$, which takes a-sides into a-sides and b-sides into b-sides. Next, let Γ_x be the vertical segment which is the intersection of the line $\mathrm{Re}\, z = x$ with $\overline{R}$ and γ_x the curve which is the image of Γ_x under f. Using the geometric obvious inequality $b' \leqslant \mathrm{length}(\gamma_x)$ and the Cauchy-Schwarz inequality one gets $K \leqslant \sup D_f$. The minimum is attained for the affine mapping. Note that the restriction to C^1-mapping is not essential. The last inequality holds for quasiconformal mapping (see, for example, [Ah2]).

B. Teichmüller approach. Teicmüller, following Grötzsch, showed that any homotopic equivalence class of quasiconformal mappings from a compact Riemann surface M to a compact Riemann surface N contains a unique mapping whose maximal dilatation K is minimum. Moreover, this unique mapping can be described geometrically in terms of holomorphic quadratic differentials on M. Such a differential gives a way of cutting up the surface M into the euclidean rectangles. These quadratic differentials locally have the form $\Phi = \varphi(z)\,dz^2$, where φ is holomorphic and they admit a picture as an orthogonal pair of foliations (horizontal and vertical lines), given locally by lines $\{\mathrm{Re}\,\zeta = \mathrm{const}\}$ and $\{\mathrm{Im}\,\zeta = \mathrm{const}\}$, where $\zeta = \int \sqrt{\varphi(z)}\,dz$, away from zeros of Φ, is natural parameter. The Teichmüller map has the form of a stretching by a factor $K^{1/2}$ along the horizontal lines in this rectangles and a shrinking by a factor $K^{-1/2}$ along vertical lines.

C. The Main Inequality. This expository paper consists of the various uniqueness theorems which follow, in general, from the length–area principle of Grötzsch. A powerful version of this principle was given by Marden and Strebel. They called it the minimum norm property for holomorphic quadratic differentials.

Marden and Strebel stated the principle by way of comparison with harmonic quadratic differentials. Gardiner gave two improvements of this principle. In the first version, one takes a minimum over all L^1-measurable quadratic differentials. These differentials satisfy an inequality of line integrals taken over arcs which are segments of regular vertical trajectories of a given quadratic differential.

In the second version, the minimum is taken over all conformal quadratic differentials satisfying an inequality of line integrals over all homotopy classes of simple closed curves. These principles lead to the following results: The Main Inequality of Reich and Strebel, the uniqueness part of Teichmüller's theorem, the sufficiency of the Hamilton-Krushkal condition for extremal dilatation.

Let Δ denote the unit disc and

$$T_\mu\varphi(z) = \frac{\left|1 - \mu(z)\frac{\varphi(z)}{|\varphi(z)|}\right|^2}{1 - |\mu(z)|^2}.$$

We will refer to the following result as the Reich–Strebel inequality or the Main Inequality.

Theorem RS (Reich and Strebel). *Suppose that f is a quasiconformal homeomorphism of Δ onto itself which is the identity on $\partial\Delta$. Then, with $\mu = \mu_f$*

$$\iint\limits_{\Delta} |\varphi(z)|\, dx\, dy \leqslant \iint\limits_{\Delta} |\varphi(z)| T_\mu\varphi(z)\, dx\, dy,$$

for every analytic integrable function φ on Δ.

Various forms of this result play a major role in the theory of quasiconformal mappings and have many applications. For applications to extremal and uniquely extremal quasiconformal mappings, we refer the interested reader to the book by Gardiner [G], and for some recent results to [MM1], [BMM], [BLMM] and [Re3].

D. The energy integral. Let M and N be two Riemann surfaces with local conformal metrics $\sigma(z)|dz|^2$ and $\rho(z)|dw|^2$ and let $f : M \mapsto N$. It is convenient for us to use the notation in local coordinates $df = (\partial f)\, dz + (\bar\partial f)\, d\bar z$ and $p = \partial f$, $q = \bar\partial f$. The energy integral (Douglas–Dirichlet functional) of f is

$$E(f, \rho) = \int_M e(f)\sigma\, dx\, dy,$$

where $e(f)$ is the energy density defined by

$$e(f)(z) = \big(|p|^2 + |q|^2\big)\frac{\rho \circ f(z)}{\sigma(z)}.$$

If ρ is the euclidean metric ($\rho = 1$ on N), then the energy integral of f is Dirichlet integral.

A critical point of the energy functional is called harmonic mapping. The Euler–Lagrange equation for the energy functional is:

$$f_{z\bar z} + (\partial(\log\rho)) \circ f\, pq = 0.$$

Thus harmonic maps arise from a geometric variational problem and as far as we know, one can not study solutions of this equation, using classical theory of elliptic equations.

In order to explain our ideas and results it is convenient to suppose that M and N are domains in $\mathbb{C}$. Recall that Δ denote the unit disc. Now, we will state a simple, but useful, property of harmonic maps (related to natural parameter). Again, we suppose, as at the beginning, that f is a harmonic mapping between Riemann surfaces M and N. Then $\varphi(z)\,dz^2$ is a holomorphic quadratic differential on M, where $\varphi = \rho \circ f p \bar{q}$ in a local coordinate.

Let P be a regular point for $\varphi(z)\,dz^2$ on M and let ζ be a natural parameter centered at P. If we compute p and q with respect to natural parameter, then we have useful formula $\rho \circ f p\bar{q} = 1$.

In Section II we will give an outline of proofs of some properties of harmonic maps, using different tools: Dirichlet's principle, minimizing sequences, different versions of Reich–Strebel inequality, etc. For general properties of harmonic maps we refer the interested reader to Eells and Lemaire ([EL1], [EL2]), Jost [J], Schoen [Sc], Schoen and Yau [SY] and further references there.

E. The Main Inequality and Dirichlet's principle. Now we will state a formula for the energy density which explains connection between Dirichlet principle for harmonic maps (in general sense) and via the Main Inequality with Grötzsch principle (an integral version of this formula appears in [ReS2], see also [We] and [M3]). Suppose that ρ is a metric density on Δ, f is C^1 function on $\overline{\Delta}$ and let h be a diffeomorphism of $\overline{\Delta}$ onto itself which is the identity on the boundary of Δ. If $\nu = \operatorname{Belt}[h]$, then

$$e(f \circ h^{-1}) = \left[\frac{1+|\nu|^2}{1-|\nu|^2} e(f) - 4\operatorname{Re}\frac{\nu}{1-|\nu|^2}\varphi\right],$$

where $\varphi = \varphi(f) = \rho \circ f\, p\bar{q}$. Hence,

$$e(f \circ h^{-1}) - e(f) = 2(|\varphi| T_\nu \varphi - |\varphi|) + r(h),$$

where

$$r(h) = r(h, f) = \frac{2|\nu|^2}{1-|\nu|^2}(e(f) - 2|\varphi|) \geqslant 0.$$

If f is a harmonic mapping (with respect to ρ), then $\varphi = \varphi(f)$ is a holomorphic function on Δ. Hence, using the Main Inequality we obtain a version of Dirichlet principle for harmonic mappings (in general sense):

$$E(f \circ h^{-1}) \geqslant E(f).$$

We expect further applications of the Main Inequality in this direction. In order to illustrate this we will outline a short proof of Dirichlet's principle (Theorem DP) for the euclidean harmonic functions using trajectories of holomorphic quadratic differentials.

Theorem DP. (Dirichtet's principle) *Suppose that*

(a) g is continuous function on $\overline{\Delta}$.

(b) g has the first partial derivatives which are continuous on Δ

(c) the energy integral of g is finite.

If u is continuous on $\overline{\Delta}$, harmonic on Δ and if $u = g$ on the boundary of Δ, then $D(g) \geqslant D(u)$, where the inequality equals if and only if $u = g$ on Δ.

Proof. Suppose that g and u are real functions and u is harmonic on $\overline{\Delta}$. There exists holomorphic function f on $\overline{\Delta}$ such that $\operatorname{Re} f = u$ on $\overline{\Delta}$. If we consider this mapping f as a natural parameter, we can divide Δ on a finite number of disjoint quadrilaterals Σ_k. In each Σ_k the mapping f is univalent and maps each Σ_k on a horizontally convex domains D_k. Using an approach as in Grötzsch principle, one can conclude

$$\iint_{\Sigma_k} |\operatorname{grad} g|^2 dx\, dy \geqslant \operatorname{area}(D_k) = \iint_{\Sigma_k} (u_x^2 + u_y^2) dx\, dy.$$

Summing these inequalities we get the Dirichlet's principle.

In Section II we state a version of Dirichlet's principle for harmonic mappings and generalize the classical area theorem in different directions (see [M1]). Also, in Section II we study uniqueness of harmonic mappings using Dirichlet's principle, minimizing sequences and different versions of the main inequality (see [MM2], [MM3] and [M1]) and we give a proof of Beurling theorem (Theorem B2) using the vertical trajectory of the corresponding holomorphic function. A review of known results in this direction is given in this section too.

F. Extremal dilatation. In this section we give a short report concerning extremal mappings. Interested reader can learn more about extremal mappings from Strebel's survey article [S6], Reich's papers [Re8], [Re9] and Earle–Li Zhong's [ELi], all of which we highly recommend. Extremal mappings have been one of the main topics in the theory of quasiconformal mappings, since its earliest days, when Grötzsch solved the extremal problem for two rectangles. In order to discuss them we need to review some familiar definitions. A homeomorphism f from a domain G onto another is called quasiconformal if f is ACL (absolutely continuous on lines) in G and $|f_{\bar z}| \leqslant k|f_z|$ a.e. in G, for some real number k, with $0 \leqslant k < 1$. It is well known that if f is a quasiconformal mapping defined on the region G, then the function f_z is nonzero a.e. in G. The function $\mu_f = f_{\bar z}/f_z$ is therefore a well defined bounded measurable function on G, called the complex dilatation or Beltrami coefficient of f. Let QC denote the space of all quasiconformal mappings from Δ onto itself. Two elements $f, g \in$QC are equivalent if $f = g$ on $\partial\Delta$. For a given $f \in$QC we denote the equivalence class of $f \in$QC by $Q_f = [f]$ or $[\mu]$, where $\mu = \mu_f$. We also use the notation $k_0([f]) = \inf\{\|\mu_g\|_\infty : g \in Q_f\}$. We let $L^\infty = L^\infty(\Delta)$ be the space of essentially bounded complex-valued measurable functions on Δ, and let M be the open unit ball in L^∞. For any μ in M there exists a quasiconformal solution $f : \Delta \mapsto \Delta$ of the Beltrami equation

$$(F1) \qquad \overline{\partial} f = \mu\, \partial f$$

unique up to a postcomposition by a Möbius transformation. We let f^μ be the solution f of (F1) normalized by $f(i) = i$, $f(1) = 1$ and $f(-1) = -1$. Two elements μ_0 and μ_1 in M are equivalent if f^{μ_0} and f^{μ_1} coincide on $\partial\Delta$. For given $\mu \in M$ the equivalence class $[\mu]$ contains at least one element μ_0 such that $\|\mu_0\|_\infty = \inf\{\|\nu\|_\infty : \nu \in [\mu]\}$. Such a μ_0 is referred to as an extremal complex dilatation and $f_0 = f^{\mu_0}$ as an extremal quasiconformal mapping (abbreviated EQC mapping).

Let Q be the Banach space consisting of holomorphic functions φ, belonging to $L^1 = L^1(\Delta)$, with norm

$$\|\varphi\| = \iint_\Delta |\varphi(z)|\,dx\,dy < \infty, \ \varphi \in Q.$$

Instead of Q the notations A and L^1_a are also used.

For $\mu \in L^\infty$ we consider the linear functional $\Lambda_\mu(\varphi) = (\mu, \varphi)$, $\varphi \in Q$, where

$$(\mu, \varphi) = \iint_\Delta \mu(z)\varphi(z)\,dx\,dy,$$

and denote by $\|\mu\|_* = \|\Lambda_\mu\|$ the norm of μ as an element of the dual space of Q. For $\mu \in L^\infty$ we say that it satisfies the Hamilton-Krushkal condition if $\|\mu\|_* = \|\mu\|_\infty$.

We are now ready to state the main result about extremal complex dilatations.

Theorem HKRS. (Hamilton–Krushkal and Reich–Strebel) *Let $\mu \in M$. A necessary and sufficient condition that f^μ is an EQC mapping is that $\|\mu\|_* = \|\mu\|_\infty$.*

G. Unique extremality. Ahlfors and Bers showed that T has a complex structure with tangent space at the base point isomorphic to Banach space Q^*. Two tangent vectors μ and ν in the tangent space to M determine the same tangent vector in T if and only if

$$\int_\Delta \varphi\mu = \int_\Delta \varphi\nu, \ \text{for all } \varphi \in Q.$$

If μ and ν have this property, we write $\mu \sim^* \nu$ and we say that they represent the same Teichmüller infinitesimal equivalence class or, more briefly, that they are infinitesimally equivalent. The space of equivalence classes is denoted by B. A given μ is said to be extremal in its infinitesimal Teichmüller class if $\|\mu\|_\infty \leqslant \|\nu\|_\infty$, for any ν infinitesimally equivalent to μ.

Recall that Hamilton, Krushkal, Reich and Strebel showed that a Beltrami coefficient ν in M is extremal in its class in T if and only if ν is extremal in its class in B. It was natural to consider whether the analogous statement holds for the unique extremality. In several articles Reich showed that in many special situations the two notions of unique extremality coincide and he conjectured that the notions may coincide in general. In [BLMM] (see also [MM1] and [BMM]) we have recently proved the answer to this conjecture is affirmative.

Theorem G1. (The Equivalence Theorem I) *μ is uniquely extremal in its Teichmüller class if and only if μ is extremal in its infinitesimally class.*

The proof of this theorem is based on estimates which allow us to compare two Beltrami coefficients μ and ν in the same global equivalence class and two Beltrami differentials in the same infinitesimal equivalence class. These estimates generalize Reich's Delta inequality for Beltrami differentials in the same equivalence class (see [R8]). Unlike Reich's forms of the Delta inequalities, our forms do not require either one of the Beltrami coefficients to have constant absolute value.

The generalized Delta inequality is our first step towards obtaining the criterion for the unique extremality of Beltrami differentials. The next important step is the analysis of the proof of Hahn-Banach theorem and its applications to our setting. In particular, we obtain the following necessary and sufficient criterion for the unique extremality of given Beltrami coefficient χ.

Theorem G2. (Characterization Theorem I) *Beltrami coefficient χ is uniquely extremal if and only if for every admissible variation η of χ there exists a sequence φ_n in $Q(\Delta)$ such that*

(a) $\delta(\varphi_n) = \|\varphi_n\| \, \|\eta\|_\infty - \operatorname{Re} \int_\Delta \varphi_n \eta \to 0$

(b) $\lim_{n\to\infty} \inf |\varphi_n(z)| > 0$, *for almost all z in $E(\eta)$.*

Here, an admissible variation η of χ is any Beltrami differential that does not increase the L^∞-norm of χ, and which is allowed to differ from χ only on the set where $|\chi(z)| \leqslant s < \|\chi\|_\infty$, where s is a constant, and the extremal set $E(\eta)$ is the set where $\eta(z) = \|\eta\|_\infty$. This criterion is analogous to the Hamilton-Krushkal, Reich–Strebel necessary and sufficient criterion for the extremality. Namely, χ is extremal if and only if there is a sequence φ_n of holomorphic quadratic differential of norm 1 such that

$$\|\chi\|_\infty - \operatorname{Re} \int_\Delta \eta\varphi_n \to 0.$$

This criterion is among listed in the theorem in Section 11, in [BLMM], which we called the Characterization Theorem. The Characterization Theorem applies to many interesting situations. For instance, we can say precisely when a Beltrami differential of the form $k|\varphi(z)|/\varphi(z)$, with φ a holomorphic quadratic differential with $\|\varphi\| = \infty$, is uniquely extremal.

There are many examples of extremal Beltrami differentials with nonconstant modulus, but all examples of uniquely extremal Beltrami differentials known up to our papers [BLMM] and [BMM] were of the general Teichmüller type. Moreover, many results obtained studying the extremal problems speak in favor of the conjecture that all uniquely extremal Beltrami differentials μ satisfy $|\mu(z)| = \|\mu\|_\infty$, for almost all z. Surprisingly, we disprove this conjecture and show that there are uniquely extremal Beltrami differentials with nonconstant modulus.

II. Dirichlet's Principle, Uniqueness of Harmonic maps and Related Problems

A. Introduction and some basic properties. The main purpose of this section is to give a short review of some results related to harmonic maps, communicated by the author and the other members of the Seminar, at University of Belgrade, during several last years. Also in this section we give a review of known results in this direction.

A1. Let M and N be two Riemann surfaces with local conformal metrics $\sigma(z)|dz|^2$ and $\rho(z)|dw|^2$ and let $f : M \mapsto N$. It is convenient for us to use notation in local coordinates $df = (\partial f)dz + (\bar\partial f)d\bar z$ and $p = \partial f$, $q = \bar\partial f$. The energy integral of f is

$$E(f, \rho) = \int_M e(f)\, \sigma \, dx\, dy,$$

where $e(f)$ is the energy density defined by

$$e(f)(z) = (|p|^2 + |q|^2)\frac{\rho \circ f(z)}{\sigma(z)}.$$

A critical point of the energy functional is called harmonic mapping. The Euler-Lagrange equation for the energy functional is:

$$f_{z\bar{z}} + \Big(\partial(\log \rho)\Big) \circ f\, pq = 0. \tag{A1}$$

Thus harmonic maps arise from a geometric variational problem and as far as we know, one can not study solutions of this equation, using classical theory of elliptic equations.

In this section we will give an outline of the proofs of some properties of harmonic maps, using different tools: Dirichlet's principle, minimizing sequences, different versions of Reich–Strebel inequality, etc. For general properties of harmonic maps we refer the interested reader to Eells and Lemaire ([EL1], [EL2]), Jost [J], Schoen [Sc], Schoen and Yau [SY] and further references there. In order to explain our ideas and results it is convenient to suppose that M and N are the domains in $\mathbb{C}$. Let Δ denote the unit disc. If $f : M \mapsto N$ is harmonic map, then $\varphi = \rho \circ f\, p\bar{q}$ is a holomorphic function. For the sake of the reader, we will sketch a proof of this result in the case when $M = \Delta$ and N is a domain in $\mathbb{C}$, with the metric $\rho(w)|dw|$.

Let λ be a complex valued function of class C^1 with compact support in Δ and let $\Phi_\epsilon(z) = z + \epsilon\lambda(z)$. Then,

$$\nu_\epsilon = \mathrm{Belt}[\Phi_\epsilon] = \frac{\epsilon\lambda_{\bar{z}}}{1 + \epsilon\lambda_z}.$$

If f is a stationary point of the energy integral, using an expression (see [ReS2]) for $E(f \circ \Phi_\epsilon^{-1}, \rho) - E(f, \rho)$, we conclude that

$$\iint_\Delta \bar{\partial}\lambda(z)\varphi(z)\, dx\, dy = 0.$$

Since φ is integrable function on Δ, it follows that φ is an analytic function on Δ, by Weyl's lemma.

Now, we will state some simple, but useful, properties of harmonic maps.

A2. Properties of harmonic maps related to natural parameter. Again, we suppose, as at the beginning, that f is harmonic mapping between Riemann surfaces M and N. Then $\varphi(z)dz^2$ is a holomorphic quadratic differential on M, where $\varphi = \rho \circ f p\,\bar{q}$ in a local coordinate.

Let P be a regular point for $\varphi(z)dz^2$ on M and let ζ be a natural parameter centered at P. If we compute p and q with respect to natural parameter, then we have important formula

$$\rho \circ f p\bar{q} = 1 \tag{A2}$$

Now, easy computation gives:

$$p\bar{q} = \frac{1}{4}\left(|f_\xi|^2 - |f_\eta|^2 - 2\,i\,\mathrm{Re}\,\bar{f}_\xi f_\eta\right)$$

Combining this formula with (A2), we find that f_ξ and f_η are orthogonal (if we consider them as vectors). Also, we can show that Jacobian $J = |p|^2 - |q|^2 = 0$ if and only if $f_\eta = 0$.

A3. Using Aronszajn's generalization of Carleman's result we can prove the following uniqueness property:

Theorem S. *If f is a harmonic mapping of an open connected set $D \subset M$ and $f = 0$ on an open subset of D, then $f = 0$ throughout D.*

General version of this result, which is concerned with the case when M and N are Riemannian manifolds, is known as Sampson's Unique Continuation Theorem (see [Sa] and [EL2]).

A4. The symmetry property.

Theorem RP. (The reflection principle) *Suppose L is a segment of the real axis, Ω^+ is a region in $H^+ = \{z : \operatorname{Im} z > 0\}$, and every $t \in L$ is the center of an open disc B_t such that $H^+ \cap B_t$ lies in Ω^+. Let Ω^- be the reflection of Ω^+: $\Omega^- = \{\bar{z} : z \in \Omega^+\}$. Suppose u is harmonic in Ω^+ and $\lim_{n\mapsto\infty} u(z_n) = 0$ for every sequence $\{z_n\}$ in Ω^+ which converges to a point on L. Then there is a function U, harmonic in $\Omega = \Omega^+ \cup L \cup \Omega^-$ such that $U = u$ in Ω^+. This function U satisfies the relation $U(z) = -U(\bar{z})$, $z \in \Omega$.*

Proof. We extend u to Ω by defining $U(z) = 0$, for $z \in L$, and $U(z) = -U(\bar{z})$, for $z \in \Omega^-$.

Example 1. It is not difficult to verify that function $f(z) = 2x + i\cos y$ is harmonic mapping from C into C with respect to the corresponding metric. This function is periodic with respect to y. The next result shows that this periodicity is typical.

Theorem M1. *Suppose that $f : \mathbb{C} \mapsto \mathbb{C}$ is a harmonic mapping, given w.r.t. natural parameter and that Jacobian of f equals zero on the real axis. Then $f(z) = f(\bar{z})$.*

The proof of this result is based on the Theorem S.

B. Dirichlet's principle and related problems. B1. If the metric density $\rho \equiv 1$ on N, then the equation (1) reduces to $f_{z\bar{z}} = 0$. In this case we say that f is a harmonic function and write $D[f]$ instead of $E(f,1)$ for the energy integral. Recall that Δ denote the unit disc. Also we will use the notation $D[\phi,\psi] = \iint_\Delta (\phi_x\psi_x + \phi_y\psi_y)\,dx\,dy$.

The following lemma is crucial in the proof of Dirichle's principle.

Lemma DP. *Suppose that*

(a) u and h are continuous on $\overline{\Delta}$ and $h \equiv 0$ on $\partial\Delta$

(b) u is harmonic on Δ and h has the continuous partial derivatives of the first order on Δ

(c) u and h have the finite Dirichlet's integral on Δ.

Then $D[u,h] = 0$.

First we will state the Dirichlet's principle for harmonic function.

Theorem DP. (Dirichtet's principle) *Suppose that*

(*a*) *g is continuous function on $\overline{\Delta}$.*

(*b*) *g has the first partial derivatives which are continuous on Δ*

(*c*) *the energy integral of g is finite.*

If u is continuous on $\overline{\Delta}$, harmonic on Δ and if $u = g$ on the boundary of Δ, then $D(g) \geqslant D(u)$, where the inequality equals if and only if $u = g$ on Δ.

Proof. If $h = g - u$, then Lemma DP shows that,

$$D[g] = D[u] + 2D[u,h] + D[h] = D[u] + D[h] > D[u],$$

unless $D[h] = 0$, i.e., h has the constant value zero.

Now, we are going to discuss some results related to Dirichlet's principle. In [M1] we gave a proof of Theorem M2 (see bellow) based on Dirichlet's principle. Before we state this result we need some definitions and we will state the area theorem and a result of Lehto–Kühnau, which motivated us.

B2. An area theorem of Lehto–Kühnau type for harmonic maps. First, we are going to prove the area theorem, which is an important tool in theory of univalent functions.

Theorem A. (The area theorem) *Let $w = f(z) = z + \frac{a_1}{z} + \cdots + \frac{a_n}{z^n} + \cdots$ be an univalent analytic function on $E = \{z : |z| > 1\}$ and let $G = \mathbb{C} \setminus f(E)$ be the omitted set. Then*

$$\pi\left(1 - \sum_{k=1}^{\infty} k|a_k|^2\right) = \text{area}(G).$$

Proof. Let K_ρ be the circle $|z| = \rho > 1$, with the positive orientation, and set

$$I_\rho = I_\rho(f) = \frac{i}{2}\int_{K_\rho} f\, d\bar{f}.$$

If $f = u + iv$ and if γ_ρ denotes the image curve of K_ρ, we have

$$I_\rho = \int_{\gamma_\rho} u\, dv$$

and by elementary calculus this represent the area enclosed by γ_ρ. Hence $I_\rho > 0$.

Direct calculation gives

$$\begin{aligned} I_\rho &= \frac{i}{2}\int_{K_\rho}\left(z + \sum_{k=1}^{\infty}\frac{a_k}{z^k}\right)\left(1 - \sum_{k=1}^{\infty} k\bar{a}_k\bar{z}^{-k-1}\right)d\bar{z} \\ &= \frac{1}{2}\int_{K_\rho}\left(z + \sum_{k=1}^{\infty}\frac{a_k}{z^k}\right)\left(\bar{z} - \sum_{k=1}^{\infty} k\bar{a}_k\bar{z}^{-k}\right)d\theta \\ &= \pi\left[\rho^2 - \sum_{k=1}^{\infty} k|a_k|^2\rho^{-2k}\right]. \end{aligned}$$

Thus $\sum_{k=1}^{\infty} k|a_k|^2\rho^{-2k} < \rho^2$, and theorem follows for $\rho \mapsto 1$.

Let us consider conformal mapping h which belongs to class Σ, i.e., h is univalent in $E = \{z : |z| > 1\}$ and has a power series expansion of the form

$$h(z) = z + \sum_{n=1}^{\infty} a_n z^{-n}$$

in E. If h has a quasiconformal extension to the plane with complex dilatation μ, satisfying the inequality $\|\mu\|_\infty = k < 1$, we say that h belongs to the subclass Σ_k of Σ. Lehto [L1], [L2] and Kühnau [K] established the area theorem for Σ_k.

Theorem LK. (Lehto–Kühnau) *Let $h \in \Sigma_k$. Then $\sum_{n=1}^{\infty} n|a_n|^2 \leqslant k^2$. The estimate is sharp.*

If we denote by P the area of the omitted set of $h(E)$, then Theorem LK states that $P \geqslant \pi(1 - k^2)$.

Before we state the Theorem M2, which is a generalization of Theorem LK to univalent harmonic mappings, we need some definitions. Let Σ' be the set of all harmonic, orientation-preserving, univalent mappings

$$h(z) = z + f(z) + \overline{g(z)} + A \log |z|$$

on E, where $f(z) = \sum_{n=1}^{\infty} a_n z^{-n}$ and $g(z) = \sum_{n=1}^{\infty} b_n z^{-n}$ are analytic on E and $A \in C$. Let Σ'_k denote the set of all homeomorphisms h of C onto itself such that: (a) the restriction of h on E belongs to Σ' and (b) the restriction of h on the unit disk $U = \{z : |z| < 1\}$ is a quasiconformal mapping with complex dilatation μ satisfying $\|\mu\|_\infty \leqslant k < 1$.

The Area theorem can be established for Σ'_k. Recall, that P denote the area of the omitted set of $h(E)$. Also, it is convenient to use notations $\tau = \sum_{n=1}^{\infty} n|a_n|^2$ and $s = 1 + 2\,\mathrm{Re}\, b_1 + l$, where $l = \sum_{n=1}^{\infty} n|b_n|^2$.

Theorem M2. *Let $h \in \Sigma'_k$. Then*
(a) $P \geqslant \pi(1 - k^2)s$; (b) The equality holds in (a) if and only if

$$h(z) = z + cz^{-1} + cg(z) + \overline{g(z)} + A \log |z|,$$

where $g(z) = \sum_{n=1}^{\infty} b_n z^{-n}$ is analytic on E; and $|c| = k$, $A \in \mathbb{C}$.

Since $P = \pi(s - \tau)$ the next result follows immediately from Theorem M1.

Corollary M1. *If $h \in \Sigma'_k$, then $\tau \leqslant k^2 s$.*

Finally we state a generalization of the area theorem to analytic functions.

Theorem A1. *Let $w = f(z) = \lambda z + \frac{a_1}{z} + \cdots + \frac{a_n}{z^n} + \cdots$ be an analytic function on $E = \{z : |z| > 1\}$ and let $G = \mathbb{C} \setminus f(E)$ be the omitted set. Then*

$$(B1) \qquad \pi\left(|\lambda|^2 - \sum_{k=1}^{\infty} k|a_k|^2\right) \leqslant \operatorname{area}(G).$$

Equality holds if and only if f is a univalent function on E.

Proof. Let K_ρ be the circle $|z| = \rho$ with positive orientation and let γ_ρ be the curve defined by the equation $w = f_\rho(e^{it}) = f(\rho e^{it}), 0 \leqslant t \leqslant 2\pi$. For given $w \neq \infty$ let $n(w)$ be the number of roots of $f(z) = w$ in $|z| > \rho$. Assume that $f \neq w$ on K_ρ and $\lambda \neq 0$. Since f has a pole of order 1 at ∞, we have $f(z) \neq w$ in $|z| \geqslant r$ for a large r and consequently, by the argument principle,

$$(B2) \qquad n(w) = \frac{1}{2\pi i}\int_{K_r - K_\rho} \frac{f'(z)}{f(z) - w}\,dz = 1 - \chi(\gamma_\rho, w),$$

where $\chi = \chi(\gamma_\rho, w)$ is the winding number (or index) of the curve γ_ρ with respect to the point w. By the analytic Green's theorem (see, for example [Po]), the area

$$(B3) \qquad I_\rho = \frac{1}{2\pi i}\int_{\gamma_\rho} \bar{w}\,dw = \frac{1}{\pi}\int_{\mathbb{R}^2} \chi(\gamma_\rho, w)\,du\,dv.$$

Let G_ρ be the set omitted by f on $E_\rho = \{|z| > \rho\}$. By (1) $w \in G_\rho$ if and only if $\chi(\gamma_\rho, w) = 1$. Also, it follows from (1) that $\chi(\gamma_\rho, w)$ is an integer less than or equal to zero if $w \notin \bar{G}_\rho$. This together with (B3) gives

$$(B4) \qquad \pi I_\rho \leqslant \operatorname{area}(G_\rho).$$

Direct calculation as in the proof of area theorem gives (B1). For the case of equality see [M].

B3. Extremal metrics and modulus. In this item we are going to give a proof of a Beurling result, which is a modification of the proof in [Ah1]. Also, we outline a new proof of the Beurling result, using minimizing sequences. Our approach is influenced by Courant's book (see [C]) and Gehring's work in $\mathbb{R}^3$ space (see [Ge1] and [Ge2]). Some generalizations of Gehring's results are presented in [AMŠ].

In unpublished work Beurling has given the following elegant and useful criterion. Before we state Beurling result we need a few definitions.

Let Ω be a region in the plane, and let Γ be family of curves and let $\rho(z) \geqslant 0$, be Borel measurable function defined in the z-plane. We say that ρ is admissible for Γ, if for every rectifiable $\gamma \in \Gamma$, $\int_\gamma \rho|dz|$ exists and $\infty \geqslant \int_\gamma \rho|dz| \geqslant 1$. In these circumstances every rectifiable arc γ has a well defined ρ-length

$$L(\gamma, \rho) = \int_\gamma \rho|dz|,$$

which may be infinite, and the open set Ω has a ρ-area $A = A_\rho = A(\rho, \Omega)$. The modulus of Γ, $M = M_\Omega(\Gamma)$, with respect to Ω, is defined as $\inf A(\Omega, \rho)$ for admissible ρ. The extremal length of Γ in Ω is defined as the reciprocal of the modulus. The extremal length is denoted by $\lambda = \lambda_\Omega(\Gamma)$.

Theorem B1. (Beurling's theorem) *The metric ρ_0 is extremal for Γ if Γ contains a subfamily Γ_0 with the following properties:*

(a) $\int_\gamma \rho_0|dz| = 1$, for all $\gamma \in \Gamma_0$;

(b) for real-valued h in Ω the conditions $\int_\gamma h|dz| > 0$ for all $\gamma \in \Gamma_0$ imply $\iint_\Omega h\rho_0\,dx\,dy \geqslant 0$.

Let Ω be an open set and let E_1, E_2 be two sets in the closure of Ω. Take Γ to be the set of connected arcs in Ω which join E_1 and E_2. The extremal length $\lambda(\Gamma)$ is called the extremal distance of E_1 and E_2 in Ω, and we denote it by $d_\Omega(E_1, E_2)$.

Example 1. The extremal distance between vertical sides of a rectangle $R = \{z = x + iy : a < x < b,\ c < y < d\}$ is $\lambda = \frac{b-a}{d-c}$.

Proof. Let $\Lambda_y = [a + iy,\ b + iy]$ and Γ_0 is the family of curves $\{\Lambda_y : c \leqslant y \leqslant d\}$. If we take $\rho_0 = 1$ Beurling's criterion is satisfied, and $\rho_0 = 1$ is extremal metric.

Example 2. Let A be the ring $A = A(r_1, r_2) = \{z : r_1 < |z| < r_2\}$. If Γ is the family of arcs in $\overline{A}$, which join circles $K_{r_1} = \{z : |z| = r_1\}$ and $K_{r_2} = \{z : |z| = r_2\}$, then

$$L(\Gamma) = \frac{1}{2\pi} \ln \frac{r_2}{r_1}. \tag{B5}$$

Proof. Let $A' = A \setminus (r_1, r_2)$ and $R = \{w : \ln r_1 < u < \ln r_2,\ 0 < v < 2\pi\}$. Since *exp* maps conformally R onto A', using the Example 1 we get (B5).

Now, we state a result of Beurling, which express the Dirichlet's integral by means of extremal distance (see [Ahl]).

Theorem B2. (Beurling's theorem) *Let Ω be a region in the complex plane bounded by a finite number of analytic Jordan curves, let E_0 and E_1 be disjoint and consist of finite number of closed arcs or curves in the boundary of Ω. Then the extremal distance $d_\Omega(E_0, E_1)$ is the reciprocal of the Dirichlet integral*

$$D(u) = \iint_\Omega (u_x^2 + u_y^2)\, dx\, dy,$$

where u satisfies:

(i) u is bounded and harmonic in Ω

(ii) u has a continuous extension to $\Omega \cup E_0^o \cup E_1^o$, and $u = 0$ on E_0 and $u = 1$ on E_1

(iii) the normal derivative $\partial u/\partial n$ exists and vanishes on C_0 (C denote the full boundary of Ω, $C_0 = C - (E_0 \cup E_1)$, and E_0^o and E_1^o denote relative interiors of E_0 and E_1 as a subset of C).

The proof of this theorem in [Ahl] is based on two important ingredients:

1) the existence of solution of a mixed Dirichlet-Neuman problem (we denote it by u)

2) decomposition of a domain on rings and quadrilateral subdomains using, in fact, the orthogonal and vertical trajectories of quadratic differential defined by u.

For the theory of trajectories of holomorphic quadratic differentials see [Ga] and [S2].

Proof of Theorem B2. Let A be the set of the end points of the E_1 and E_2 as subsets of C. The reflection principle shows that u has a harmonic extension across $\partial\Omega \setminus A$.

Let $z_0 \in A$, for example, $z_0 \in E_1$. We can chose a local conjugate v in Ω near z_0 such that, on the boundary, $u = 0$ on one side of z_0 and $v = 0$ on the other side of z_0. Then, by the reflection principle, there exists neighborhood V of z_0 and an analytic function φ in $V \setminus \{z_0\}$ such that $\varphi = (u + iv)^2$ in $\Omega \cap V$. Hence, φ is an analytic function on V and has a simple zero at z_0. Therefore, $u_x - iu_y$ must tend to ∞, and the number of critical points in $\overline{\Omega} \setminus A$ is finite.

Locally, for every $z_0 \in \partial\Omega \setminus A$ there exists a neighborhood V of z_0 and an analytic function f on V such that $\operatorname{Re} f = u$ on V. Hence, we can define horizontal trajectories with respect to $w = f(z)$.

The part of noncritical horizontal trajectory γ which is in $\overline{\Omega}$ can be parameterized with parameter interval $I = [0, 1]$ such that:

1. γ join E_1 and E_2 in Ω (more precisely $\gamma(0,1) \subset \Omega$, $\gamma(0) \in E_1$ and $\gamma(1) \in E_2$).
2. $\operatorname{Re}\gamma$ is strictly increasing function on I and $\operatorname{Re}\gamma(0) = 0$, $\operatorname{Re}\gamma(1) = 1$.

Hence, we conclude that up to a set of Lebesgue 2-dimensional measure zero there exists finite number of disjoint quadrilateral Σ_k, $k = 1, 2, \ldots, n$, such that:

1. $\Omega = \bigcup_{k=1}^{n} \Sigma_k$
2. Each Σ_k is swept out with noncritical horizontal trajectories
3. There exists rectangles R_k of width 1 and height m_k and conformal (univalent) mapping $\Phi = \Phi_k$ of Σ_k onto R_k such that $\operatorname{Re}\Phi_k = u$ on Σ_k. Hence,

$$m_k = \iint_{\Sigma_k} |\Phi'|^2 dx\, dy \quad \text{and} \quad m = \sum_{k=1}^{n} m_k = D(u).$$

Together rectangles R_k fill out a rectangle with sides 1 and $D(u)$. After appropriate identification we obtain a model of Ω with E_1 and E_2 as vertical sides.

From this model and Beurling theorem (Theorem B1) it is immediately clear that the euclidean metric is extremal, and we conclude that $d_\Omega(E_1, E_2) = 1/D(u)$.

Our first purpose was to give more elementary proof of this result (that is, with no use of these two subjects), using a minimizing sequence (see, for example Courant's book [C]), and to derive some equalities not contained in the proof of Beurling's theorem. During our work on this problem we become aware of Gehring's works (see [Ge1] and [Ge2]), which strongly influenced our research.

In [Ge1] and [Ge2] Gehring proved that essentially Väisälä's definition of extremal distance between E_0 and E_1 in Ω is equivalent to the Dirichlet's integral definition due to Loewner (see [Lo]) if Ω is a ring domain in $\mathbb{R}^3$, and E_0 and E_1 are boundary components of Ω. Gehring used this result to study quasiconformal mappings in space. We generalize this result to the setting of smooth domains in $\mathbb{R}^n$. An application of this result gives a short proof of Beurling's Theorem. As we understand, there are additional technical difficulties if we work with general domains instead of ring domains. Before we state the result we need a few definitions.

Definition B1. Let Ω be an open set in $\mathbb{R}^n$ and Γ a set whose elements γ are rectifiable arcs in Ω. Let ρ be a nonnegative Borel measurable function in Ω (such ρ we will call metric). We can define the ρ-length of γ by

$$L(\gamma, \rho) = \int_\gamma \rho\, |dx|$$

the ρ-volume of Ω as

$$V(\Omega,\rho)=\int_\Omega \rho^n\, dV(x)$$

where dV is the n-dimensional Lebesgue measure in $\mathbb{R}^n$, and the minimum length of Γ by $L(\Gamma,\rho)=\inf_{\gamma\in\Gamma} L(\gamma,\rho)$. The modulus of Γ in Ω is defined by

$$\operatorname{mod}_\Omega(\Gamma)=\inf_\rho \frac{V(\Omega,\rho)}{L(\Gamma,\rho)^n},$$

where ρ is subject to the condition $0<V(\Omega,\rho)<\infty$. The extremal length of Γ in Ω is defined as $\Lambda_\Omega(\Gamma)=\ \operatorname{mod}_\Omega(\Gamma)^{1/1-n}$.

Definition B2. Let Ω be an open set in $\mathbb{R}^n$, and let E_0, E_1 be two sets in the closure of Ω. Take Γ to be the set of connected arcs in Ω which join E_0 and E_1, i.e. each $\gamma\in\Gamma$ has one endpoint in E_0 and the other in E_1 and all other points of γ are in Ω. The extremal length $\Lambda(\Gamma)$ is called the extremal distance of E_0 and E_1 in Ω, and we denote it by $d_\Omega(E_0,E_1)$.

Now, let Ω be a bounded region whose boundary consists of a finite number of C^1 hypersurfaces. If E_0 and E_1 are disjoint, and each is a finite union of closed hypersurfaces contained in the boundary of Ω, then we define the conformal n-capacity of Ω as

$$C[\Omega,E_0,E_1]=\inf_u \int_\Omega |\nabla u|^n\, dV(x),$$

where infimum is taken over all functions $u:\Omega\to\mathbb{R}$ which are differentiable in Ω, continuous in $\overline{\Omega}$ and have boundary values 0 on E_0 and 1 on E_1.

The proof of the following theorem is given in [AMŠ].

Theorem AMŠ. *If Ω is a bounded domain, whose boundary consists of a finite number of C^1 hypersurfaces, and if E_0 and E_1 are disjoint sets of the boundary of Ω consisting of finite number of closed hypersurfaces, then we have*

$$\operatorname{mod}_\Omega(\Gamma)=\inf_f \frac{V(\Omega,f)}{L(\Gamma,f)^n}=C[\Omega,E_0,E_1],$$

where f is any metric in Ω and Γ is the family of all Jordan arcs joining E_0 and E_1 inside Ω.

The case $n=2$ of previous theorem enables us to give a short proof of Theorem B. In fact, the proof immediately follows from Theorem 1.3 (see [C]), which gives a solution of a mixed Dirichlet-Neuman problem.

The proof of Theorem 1.3 in Courant's book [C], is based on using minimizing sequences. We believe that we can use minimizing sequences as Gehring in [Ge1] to show existence of the extremal admissible function $u\in E(\Omega,E_0,E_1)$ such that

$$C[\Omega,E_0,E_1]=\int_\Omega |\nabla u|^n dV.$$

B4. Dirichlet's principle for harmonic mappings. Let N be complete Riemannian manifold of dimension n and let its metric in local coordinates be given by (g_{ik}), with Christoffel symbols Γ^i_{kl}. For $f \in H^{1,2}(M,N)$ we define the energy density

$$e(f)(z) = \frac{1}{\sigma^2}\sum g_{ik}(f)\,(f^i_x f^k_x + f^i_y f^k_y),$$

and the energy as

$$E(f) = \frac{1}{2}\int_M e(f)\sigma^2 dx\,dy = \frac{1}{2}\int_M \sum g_{ik}(f),(f^i_x f^k_x + f^i_y f^k_y)\,dx\,dy,$$

where we write $f = (f^1, f^2, \dots, f^n)$ in local coordinates. A solution of the corresponding Euler–Lagrange equation $\Delta f^i + \Gamma^i_{kl}(f^k_x f^l_x + f^k_y f^l_y) = 0$, $i = 1, 2, \dots, n$, is called harmonic map.

Theorem M3. (Dirichlet's principle for harmonic mappings) *Let N be Riemannian n-dimensional manifold and $f : \overline{\Delta} \mapsto N$ be a harmonic mapping. If Φ is diffeomorphism of $\overline{\Delta}$ onto itself, which is identity on $\partial\Delta$, then $E(f \circ \Phi) \geqslant E(f)$.*

C. Uniqueness of harmonic maps. Our further discussion is concerned mainly with the case when M and N are domains in complex plane $\mathbb{C}$. Recall, that the following result enables us to use theory of trajectory of holomorphic quadratic differentials.

C1. If f is a harmonic mapping between Riemann surfaces M and N with local conformal metrics $\sigma(z)|dz|^2$ and $\rho(w)|dw|^2$, respectively, then $\varphi = \rho p \bar{q}\, dz^2$ is a holomorphic quadratic differential. For example if M and N are subset of the complex plane C, this simply means that the function $\rho p \bar{q}$ is a holomorphic function. This enables us to use the techniques and results from the theory of holomorphic functions.

C2. Marković and the author, using a version of Reich–Strebel inequality, proved the following uniqueness property.

Theorem MM. *Suppose that*

(*a*) *f and g are harmonic diffeomorphisms of Δ onto itself*

(*b*) *f and g are continuous on $\overline{\Delta}$*

(*c*) *$f = g$ on $\partial\Delta$.*

If, in addition, we suppose that the energy integrals of f and g are finite, they are identical.

This result was communicated on our Seminar at Belgrade University in 1996. and at Nevannlina Colloquium, Switzerland 1997. The proof is based on the next lemma if f and g are diffeomorphisms of $\overline{\Delta}$ onto itself and on a new version of Reich–Strebel inequality in general case.

Lemma MM. *Suppose that f and g are diffeomorphisms of $\overline{\Delta}$ onto $\bar{\Delta}$ and that f is harmonic with respect to conformal metric $ds = \rho(w)|dw|$ on $\bar{\Delta}$. If we suppose in addition, that $E(f) < +\infty$ and that $f = g$ on $\partial\Delta$, then*

$$\int_\Delta \tilde{\rho}(\zeta)\,d\xi\,d\eta \leqslant \int_\Delta \tilde{\rho}(\zeta)\frac{1-|\tilde{\mu}(\zeta)|}{1+|\tilde{\mu}(\zeta)|}\frac{\left|1+\frac{\tilde{\chi}(\zeta)}{\tilde{\mu}(\zeta)}|\tilde{\mu}(\zeta)|\right|^2}{1-|\tilde{\chi}(\zeta)|^2}\,d\xi\,d\eta,$$

where $\tilde{\mu} = \mathrm{Belt}(f^{-1})$, $\tilde{\chi} = \mathrm{Belt}(g^{-1})$ *and* $\tilde{\rho}(\zeta) = \rho(\zeta)\dfrac{|\tilde{\mu}(\zeta)|}{1-|\tilde{\mu}(\zeta)|^2}$.

We will outline a proof of Theorem MM in the case that $f = Id$ on $\partial\Delta$ and that f is diffeomorphism of $\overline{\Delta}$ onto itself. For the proof it is useful to observe that if f is harmonic, then Beltrami dilatation μ of f has the form

$$\mu(z) = s(z)\frac{|\varphi(z)|}{\varphi(z)},$$

where s is non-negative measurable function and $\varphi = \rho \circ f p\bar{q}$ is an analytic function. Thus we have that expression $\mu\varphi/|\varphi|$, which appears in Reich–Strebel inequality equals $|\mu|$ and we get

$$\int_\Delta |\varphi|\,dx\,dy \leqslant \int_\Delta |\varphi|\frac{1-|\mu|}{1+|\mu|}\,dx\,dy.$$

If φ is not identically zero we get $\mu = 0$ a.e. Hence we conclude that f is conformal mapping. Since $f = \mathrm{Id}$ on $\partial\Delta$, we get that $f = \mathrm{Id}$ on Δ. In general, we need a version of main inequality which holds for the mapping whose maximal dilatation can be 1.

C3. Marković and the author have proved that $f = g$ under weaker conditions, then in Theorem MM. The following two results will appear in [MM3].

Theorem MM1. *Suppose that*
(*a*) f *is homeomorphism of* $\overline{\Delta}$ *onto itself*
(*b*) f *has the first generalized derivatives on* Δ
(*c*) f *is identity on* $\partial\Delta$
(*d*) f *is harmonic w.r.t. some metric density* ρ *on* Δ
(*e*) *Hopf differential of* f *is integrable on* Δ.
Then f *is the identity on* Δ.

Theorem MM2. (The uniqueness property). *Suppose that*
(*a*) f *and* g *are homeomorphisms of* $\overline{\Delta}$ *onto itself and* $f = g$ *on* $\partial\Delta$
(*b*) f *and* g *are loc. q.c. on* Δ
(*c*) f *and* g *are harmonic w.r.t. some metric density* ρ *on* Δ
(*d*) *Hopf differentials of* f *and* g *are integrable on* Δ.
Then f *and* g *are identical.*

Also, we might add that we have a generalization of this result if instead of the unit disk, we consider Riemann surfaces. Recall, if the metric $\rho \equiv 1$ on N, which is open subset of complex plane $\mathbb{C}$ (euclidean case), we will say harmonic function instead of harmonic mapping. Thus in euclidean case this result says that the solution of classical Dirichlet problem is unique.

The proof of Theorem MM2 is based on a new version of Riech-Strebel inequality. Note that if f and g are harmonic property (A) says that function $\varphi = \rho \circ f p\bar{q}$ and $\psi = \rho \circ g A\bar{B}$ are holomorphic functions on the unit disk, where we use notation $A = \partial g$, $B = \bar{\partial} g$. The idea of the proof is to apply a new version of Reich–Strebel inequality to functions φ and ψ.

In the next item we are going to give a short discussion of known result related to uniqueness of harmonic maps.

C4. We refer the interested reader to [J] for the global uniqueness theorem of Al'ber and Hartman, for the result of Jäger and Kaul and for further references.

Theorem AH (Al'ber and Hartman). *Let $u : M \mapsto N$ be a harmonic map between compact Riemannian manifolds (without boundary). Suppose N has negative sectional curvature. Then u is unique harmonic map in its homotopy class unless $u(M)$ is a point or a closed geodesic.*

If the sectional curvature of N is non-positive, then for any two homotopic harmonic $u_0, u_1 : M \mapsto N$, there exist a family $u_t : M \mapsto N$ of harmonic maps, with the property that the curves $u_t(x)$, for fixed $x \in M$, $t \in [0,1]$ varying, constitute a family of parallel geodesics, parameterized proportionally to arc length. In particular, all maps u_t have the same energy.

Theorem JK (Jäger and Kaul). *Suppose that $u_i : \overline{\Omega} \mapsto N$ $(i = 1,2)$ are harmonic maps of class $C^0(\overline{\Omega}, N) \cap C^2(\Omega, N)$, Ω is a bounded domain in some Riemannian manifold, and $u_i(\overline{\Omega}) \subset B(p,\rho)$, where $B(p,\rho)$ is a geodesic ball in N, disjoint to the cut locus of p and with radius $\rho < \pi/2\kappa$, where κ^2 is an upper bound for the sectional curvature of $B(p,\rho)$. If $u_1 = u_2$ on $\partial\Omega$, then $u_1 \equiv u_2$.*

We refer the interested reader to the Schoen-Yau book [SY] for uniqueness theorems concerning harmonic maps into non-positive curved metric spaces and further references.

After writing the previous version E. Reich pointed out to us that H. Wei [We] studied uniqueness property of harmonic mappings. Also, we became aware of the Coron–Helein paper [CH]. H. Wei using the formula for the energy of variation of a mapping (see [ReS2]) and Reich–Strebel inequality, proved a weaker version of Theorem 2 concerning q.c. mapping. Namely, H. Wei proved Theorem MM2 under additional hypotheses that

(c) f and g are q.c. mappings on the unit disk Δ onto itself

(d) the metric density ρ is an integrable function on Δ.

Note that the hypotheses (c) and (d) provide that the energy integral of f and g are finite.

In [CH], Coron–Helein used completely different approach then H. Wei in [We] to study minimizing harmonic mappings. Their approach was based on decomposition of given metric g on Δ as the sum of two metrics c and h such that c is conformal metric of the euclidean metric e, h has non-positive Gaussian curvature and Id is harmonic map between (Δ, e) and (Δ, h).

Theorem CH (Coron–Helein). *Let (M,h) and (N,g) be two Riemannian compact surfaces of class C^∞ possibly with boundary. Then any smooth harmonic diffeomorphism between (M,h) and (N,g) is minimizing in its homotopy class. Moreover, if ∂M is nonempty or if the genus of M is strictly larger then one, then such a diffeomorphism is the unique minimizing map in its homotopy class.*

D. Related results. First, we will give an application of Theorem MM2 in the case when the energy integral is infinite.

D1. Suppose that

(a) f and g are harmonic diffeomorphisms from the Δ onto itself w.r.t. Poincaré metric.

(b) Hopf differentials $\varphi = \mathrm{Hopf}(f)$ and $\psi = \mathrm{Hopf}(g)$ are integrable on Δ.

Since φ and ψ belong to Bers space (see, for example, [Ah2], [W] and [AMM] for definition and properties of Bers space) a result of Wan [W] shows that f and g are q.c. mappings of Δ onto itself. If, in addition, we suppose that $f = g$ on the boundary of the unit disk, an application of Theorem MM2 shows that f and g are identical.

Note that every harmonic diffeomorphism of Δ onto itself w.r.t. Poincaré metric has infinite energy integral.

The following example shows that without assumption that Hopf differentials are integrable Theorem MM1 is not valid.

D2. Let φ be the conformal mapping of the unit disk Δ onto upper half-plane H and let $\rho(w) = |\varphi'(w)|$. Next, let $g = \psi \circ h \circ \varphi$, where ψ is the inverse function of φ and h is given by $h(z) = x + iky$, $k > 0$. We leave to the reader to verify that g is q.c. harmonic mapping (w.r.t. ρ) of the unit disk Δ onto itself and that $g = \mathrm{Id}$ on the boundary of Δ.

Although, the metric defined by the density ρ is flat on the complex plane C except at one point, Theorem MM1 is not valid.

D3. In connection with the parts (D1) and (D2) of this section, we will give a short discussion (we follow Schoen [Sc]).

There is an interesting conjecture which is due to Schoen (see also [Sc]).

Conjecture. The q.c. harmonic homeomorphisms from the unit disk Δ onto itself, w.r.t. Poincaré metric, are parameterized by the boundary values of q.c. maps of the disk.

This is a question which involves proving both an existence and a uniqueness theorem. The existence result for this ideal boundary value problem has been shown by Li and Tam [LT1] under the additional hypothesis that boundary map be sufficiently differentiable. They have also obtained counterexamples to uniqueness without the quasi-conformal hypothesis (but with continuity) and then proved the uniqueness part of Schoen's conjecture (see [LT2]).

A result of Wan [W] gives a parameterisation of the q.c. harmonic homeomorphisms of Δ in terms of bounded holomorphic quadratic differentials on Δ. Wan has shown that if f is q.c. mapping, then Hopf differential of f is bounded w.r.t. the Poincaré metric on Δ. Conversely, he has shown that for any bounded holomorphic quadratic differential Φ on Δ there is a unique q.c. harmonic homeomorphism $f : \Delta \mapsto \Delta$ such that $\mathrm{Hopf}(f) = \Phi$.

D4. Theorem MM2 remains valid if the condition (b) (in the hypotheses of Theorem MM2 is replaced by the following.

(e) f, g and their inverse mapping have L^2-derivatives.

The idea of the proof is as follows. If the condition (e) holds then one can get that $f \circ g^{-1}$ and $g \circ f^{-1}$ have L^1-derivatives and its partial derivatives satisfy the chain rule (for a details see Lemma 6.4 of [LV, p. 151]).

It is well-known that the condition (b) implies the condition (e) (see, for example, [LV]).

For a development of theory of harmonic mappings by means of Sobolev spaces, we refer to Schoen-Yau book [SY].

D5. Harmonic maps and extremal QC mapping. Before we state the results, we need some notations. Suppose that f is quasi-conformal mapping of the unit disk Δ onto itself. Let $k[f] = \operatorname{ess\,sup}\{|\mu_f(z)| : z \in \Delta\}$ and let $Q(f)$ denote the collection of all q.c. mappings of Δ whose pointwise boundary values on $\partial\Delta$ agree with those of f. We call f extremal (in its Teichmüller class) if $k[f] \leqslant k[g]$ for every $g \in Q(f)$. An extremal q.c. mapping f is uniquely extremal (in its Teichmüller class) if $k[f] < k[g]$ for every other g in $Q(f)$.

Theorem M4. (The first removable singularity theorem). *Suppose that*

(a) f is q.c. mapping from Δ onto Δ

(b) f is a harmonic function with respect to the metric density ρ on $\Delta \setminus K$, where K is compact subset of Δ

(c) f is extremal in its Teichmüller class

(d) there are two positive constant m and M such that $m \leqslant |\varphi(z)| \leqslant M$ for each $z \in \Delta \setminus K$, where φ is Hopf differential of f.

Then φ has an analytic extension $\tilde{\varphi}$ from $\Delta \setminus K$ to Δ and $\mu(z) = k\,|\tilde{\varphi}(z)|/\tilde{\varphi}(z)$ a.e. in Δ, where k is a constant.

Theorem M5. (The second removable singularity theorem). *Suppose that*

(a) f is uniquely extremal q.c. mapping, in its class, from Δ onto Δ

(b) f is a harmonic function with respect to the metric density ρ on $\Delta \setminus K$, where K is compact subset of Δ.

Then we have the same conclusion as in the previous theorem.

During our work with Božin on the problems related to uniquely extremal q.c. mapping [BMM], we also obtained some results of this type.

III. New version of the Main Inequality

Analyzing the proof of the Grötzsch principle we discovered the following lemma. Let D be a vertically convex domain of finite area in the complex plane $\mathbb{C}$ and let F be a mapping from the domain D onto the domain G. Suppose that we have metric $ds = \rho(w)|dw|$ on G. Let Γ_x be the interval which is the intersection of D by the straight line $\operatorname{Re} z = x$ and let γ_x be the curve which is the image of Γ_x under F. Let $p(x,y) = x$ be the projection and let $(\alpha,\beta) = p(D)$.

Lemma 1. *With the notation and hypothesis just stated, suppose (in addition) that the mapping F is homeomorphism which has the first generalized derivatives and that*

$$\operatorname{length}(\Gamma_x) \leqslant \int_{\gamma_x} \rho(w)|dw| \quad \textit{a.e. in } (\alpha,\beta).$$

Then

$$\operatorname{area}(D) \leqslant \left[\iint_G \rho^2(w)\,du\,dv\right]^{1/2}\left[\iint_D T_\nu\,d\xi\,d\eta\right]^{1/2},$$

where $\nu = \mathrm{Belt}[F]$.

The proof of this lemma will be given in the section C of this section. Note that this lemma enables us to get a new version of the Main Inequality (Theorem 1, section D) which is applicable to mappings which are not quasiconformal. But first, we are going to give a proof of the Reich–Strebel inequality in the case of Riemann surfaces of finite analytic type using this lemma.

A. The inequality of Reich and Strebel. It is convenient to use notation

$$T_\mu\varphi(z) = \frac{\left|1-\mu(z)\frac{\varphi(z)}{|\varphi(z)|}\right|^2}{1-|\mu(z)|^2}.$$

Theorem RS. (Reich and Strebel) *Let R be a Riemann surface of finite analytic type and let φ be integrable holomorphic quadratic differential on R. Let f be a quasiconformal self mapping of R which is homotopic to the identity map and let μ be the Beltrami coefficient of f. Then*

$$\iint_R |\varphi(z)|\,dx\,dy \leqslant \iint_R |\varphi(z)| T_\mu\varphi(z)\,dx\,dy. \tag{A1}$$

Proof. Suppose that all noncritical trajectory of φ are closed. Up to a set of Lebesgue 2-dimensional measure zero $R = \bigcup \Sigma_k$, where Σ_k are disjoint ring domains. Each Σ_k is swept out by a family of vertical trajectories of the holomorphic quadratic differential $\varphi(z)dz^2$, and in each Σ_k there exists a single valued univalent branch $\zeta = \Phi_k(z)$ of $\int \sqrt{\varphi(z)}\,dz$.

Each region $\Phi_k(\Sigma_k)$ is a rectangle $R_k = \{\zeta : 0 < \xi < a_k,\ 0 < \eta < b_k\}$. Let $\Psi_k = \Phi_k^{-1}$, $F_k = f \circ \Psi_k$, $\theta = \theta_\xi = \Psi_k(\Gamma_\xi)$, where Γ_ξ is vertical interval which is the intersection of R_k with the straight line $\operatorname{Re}\zeta = \xi$, and let $\gamma = \gamma_\xi = f(\theta_\xi)$. Since the closed trajectories are shortest in their homotopy class (see Theorem 17.1 of [S2]), we obtain

$$b_k = \int_\theta \sqrt{|\varphi(z)|}|dz| \leqslant \int_\gamma \sqrt{|\varphi(w)|}|dw|.$$

If $G_k = f(\Sigma_k)$, $\nu = \nu_k = \mathrm{Belt}[F_k]$,

$$A_k = \left[\iint_{G_k} |\varphi(w)|\,du\,dv\right]^{1/2} \quad\text{and}\quad B_k = \left[\iint_{D_k} T_\nu\,d\xi\,d\eta\right]^{1/2},$$

then by Lemma 1 $a_k\, b_k \leqslant A_k\, B_k$. Using the change of variables $z = \Phi_k^{-1}(\zeta)$, we get

$$B_k{}^2 = \iint_{\Sigma_k} |\varphi(z)|\; T_\mu\varphi(z)\,dx\,dy.$$

Now, an application of Cauchy-Schwarz inequality, as in the proof of the new version of the Main Inequality (see bellow), gives (A1).

Using the fact that quadratic differentials with closed trajectories are dense in Q (see Theorem 25.2 of [S2]), we can get the Main Inequality in general.

B. The Problem of Grötzsch. B1. In order to motivate the statement and proof of our version of the Main Inequality we will emphasize the main points in the proof of Grötzsch's principle. We will follow [MM2], where we announced the results of this section.

If Q is a square and R is a rectangle, not a square, there is no conformal mapping of Q on R which maps vertices on vertices. Instead, Grötzsch asked for the most nearly conformal mapping of this kind and took the first step toward the creation of a theory of quasiconformal mappings.

Let $w = f(z)$ be a mapping from one region to another. Recall that we use notation $df = p\,dz + q\,d\bar{z}$, where $p = \partial f$ and $q = \bar{\partial} f$. The complex (Beltrami) dilatataion is $\mu_f = \text{Belt}[f] = q/p$. The dilatation of f is:

$$D_f = \frac{|p| + |q|}{|p| - |q|}.$$

We pass to the Grötzsch problem and give it a precise meaning by saying that f is most nearly conformal if $\sup D_f$ is as small as possible. Let R, R' be two rectangles with sides a, b and a', b'. We may assume that $K = \frac{b'}{a'} : \frac{b}{a} \geqslant 1$. The mapping f is supposed to be C^1-homeomorphism from $\overline{R}$ onto $\overline{R'}$, which takes a-sides into a-sides and b-sides into b-sides. Next, let Γ_x be the vertical segment which is the intersection of the line $\operatorname{Re} z = x$ with $\overline{R}$ and γ_x the curve which is image of Γ_x under f.

The starting point of Grötzsch's approach is the geometric obvious inequality

$$(B1) \qquad b' \leqslant \text{length}(\gamma_x) = \int_0^b |p - q|\, dy.$$

Using:

$$(B2) \qquad \iint_R J_f \, dx\, dy = a'b',$$

where J_f denotes the Jacobian of f, and the Cauchy-Schwarz inequality one gets

$$(B3) \qquad K \leqslant \sup D_f.$$

The minimum is attained for the affine mapping. We note the following connections to with Grötzsch's problem. The restriction to C^1-mapping is not essential. The inequality (B3) holds for quasiconformal mapping (see, for example, [Ah2]).

In order to give a version of Grötzsch principle concerning mappings with L^1-derivatives, we need the following definition.

B2. Definition of L^p-derivatives. Let D be a domain in C. We say that a function $f : D \mapsto C$ has L^p-derivatives, $p \geqslant 1$, if it satisfies the following two conditions:

(a) f is absolutely continuous on lines in D

(b) The partial derivatives f_x and f_y belong to L^p in every compact subset of D.

When we say that f has first generalized derivatives in D this means that f has L^1-derivatives in D. For various characterizations of functions with L^p-derivatives and their important role in the theory of quasiconformal mappings we refer to Chapters III to VI of the book by Lehto Virtanen [LV].

B3. A version of Grötzsch's principle. Before, we give further extension of Grötzsch's principle, it is useful to consider the following example when (B1) and (B3) does not hold.

Example 1. Let $\alpha : I \mapsto I$, where $I = [0,1]$, be Cantor function and let $f(z) = x + i(y + \alpha(y))$.

Note that this function does not satisfy ACL property and that the known formula for the length of curve by means of first partial derivatives does not hold.

Suppose that

(a) f is a homeomorphism of closed rectangle $\overline{R}$ onto the closed rectangle $\overline{R'}$ which maps a-sides onto a'-sides and b-sides onto b'-sides.

(b) f has the first generalized derivatives on R.

In order to get some conclusion we can follow the outline of the proof of Grötzsch principle from the subsection B1 of this section. We need the following definition. At the point z where $\mu(z)$ is defined and $|\mu(z)| \neq 1$ we define $T_\mu(z)$ by

$$T_\mu(z) = \frac{|1-\mu(z)|^2}{1-|\mu(z)|^2}$$

Also, at point z where $|p(z)| = |q(z)|$ we define $T_\mu(z)$ to be zero if $p(z) = q(z)$ and $+\infty$ if $p(z) \neq q(z)$.

Now, we can give the precise meaning of $T_\mu\varphi$ by means of T_χ, where $\chi = \mu\varphi/|\varphi|$.

Since f satisfies the ACL-property, inequality (B1) holds for a.a. $x \in [0,a]$.

In order to prove inequality (B4) (see below), we can suppose that T_μ is defined and finite a.e. on R, because otherwise the right-hand side of (B4) is infinite.

Next, we can integrate w.r.t. dx over $[0,a]$ and use the fact that the Jacobian

$$J_f = |p|^2(1-|\mu|^2) \quad \text{a.e. on } R.$$

Instead of (B2), we have

$$\int_R J_f\, dx\, dy \leqslant \operatorname{area}(R') = a'b'.$$

Now, an application of Cauchy-Schwarz inequality gives

$$\text{area}(R)\frac{b'}{b} \leqslant [\text{area}(R')]^{1/2}\left[\iint_R T_\mu\right]^{1/2}. \tag{B4}$$

Further development of the ideas outlined above leads us to Lemma 1 (see below), which will be used in the proof of the new version of the Main Inequality.

C. Proofs of Lemma 1 and Lemma 2. For the sake of the reader recall the statement of the Lemma 1.

Let D be a vertically convex domain of finite area in the complex plane $\mathbb{C}$ and let F be a mapping from the domain D onto the domain G. Suppose that we have metric $ds = \rho(w)|dw|$ on G. Let Γ_x be the interval which is the intersection of D by the straight line $\operatorname{Re} z = x$ and let γ_x be the curve which is the image of Γ_x under F. Let $p(x,y) = x$ be the projection and let $(\alpha,\beta) = p(D)$.

Lemma 1. *With the notation and hypothesis just stated, suppose (in addition) that the mapping F is homeomorphism which has first generalized derivatives and that*

$$\text{length}(\Gamma_x) \leqslant \int_{\gamma_x} \rho(w)|dw| \quad \text{a.e. in } (\alpha,\beta). \tag{C1}$$

Then

$$\text{area}(D) \leqslant \left[\iint_G \rho^2(w)\,du\,dv\right]^{1/2}\left[\iint_D T_\nu\,d\xi\,d\eta\right]^{1/2}, \tag{C2}$$

where $\nu = \text{Belt}[F]$.

Proof. We will use the notation $dF = P\,d\zeta + Q\,d\bar\zeta$, where $P = \partial F$ and $Q = \bar\partial F$. We can suppose that T_ν is defined and finite a.e. on D, because otherwise the right-hand side of (C2) is infinite. With definition of T_ν in mind, this means that $P = Q$ a.e. on A, where A is the set on which Jacobian J_F equals zero. Since F is absolutely continuous on Γ_x for a.a. $x \in (\alpha,\beta)$, we find

$$\rho - \text{length}(\gamma_x) = \int_{\Gamma_x} (\rho\circ F)(\zeta)|P||1-\nu|\,d\eta.$$

By Fubini's theorem and assumption (C1),

$$(\text{area})(D) \leqslant \iint_D (\rho\circ F)(\zeta)|P||1-\nu|\,d\xi\,d\eta.$$

Since $J_F = |P|^2(1-|\nu|^2)$ a.e. on D, the term on the right can be written in the form

$$\tau = \iint_D (\rho\circ F)(\zeta)J_F^{1/2}T_\nu^{1/2}\,d\xi\,d\eta.$$

Next, using the Cauchy-Schwarz inequality we conclude that $\tau \leqslant A^{1/2} \cdot B^{1/2}$, where

$$A = \iint_D (\rho^2 \circ F)(z)\, J_F(z)\, dx\, dy \quad \text{and} \quad B = \left[\iint_D T_\nu \, d\xi\, d\eta\right].$$

Let $C = \iint_G \rho^2(w)\, dudv$. We need the following lemma to finish the proof.

Lemma 2. *We have $A \leqslant C$.*

Proof. Let the measure μ be defined by

$$\mu(E) = \int_E J_F(z)\, dx\, dy \quad \text{and} \quad \mu_F(E) = m(F(E)),$$

for every Lebesgue measurable set E. Since F is a homeomorphism which possesses finite partial derivatives a.e. in D, by Lemma 3.3 of [LV, p. 131], we have $\mu(E) \leqslant \mu_F(E)$, and therefore we have the desired result.

D. Proof of new version of the Main Inequality. There are several papers of Reich and Strebel which concern various forms of the Main Inequality. Our proof is based on their ideas.

Here, we will give a complete proof of the Theorem 1, because we need to be careful when we work with mappings whose dilatation is not bounded. For convenience of the reader let us recall the statement of Theorem 1.

Theorem 1. *Suppose that*
(*a*) *f is a homeomorphism of $\overline{\Delta}$ onto itself*
(*b*) *f has first generalized derivatives on Δ*
(*c*) *f is the identity on $\partial\Delta$.*
Then the inequality

$$\iint_\Delta |\varphi(z)|\, dx\, dy \leqslant \iint_\Delta |\varphi(z)| T_\mu \varphi(z)\, dx\, dy$$

holds for every analytic integrable function φ on Δ.

D1. First, we observe that Theorem 1 can be reduced to the case when φ is also analytic on $\partial\Delta$. Namely, let φ_r, $0 < r < 1$, be the function defined by $\varphi_r(z) = \varphi(rz)$, $z \in \Delta$. If Theorem 1 holds for every φ_r, $0 < r < 1$, then, when r approaches 1 we conclude that the theorem holds for φ, by Lebesgue dominated convergence theorem.

Suppose that φ is an analytic function in $\overline{\Delta}$. The following decomposition is possible (see [S1] and [S5]). Up to a set of Lebesgue 2-dimensional measure zero, $\Delta = \bigcup_{k=1}^{n} \Sigma_k$, where $\{\Sigma_k\}$ are disjoint simple connected "strip" domains. Each Σ_k is swept out by a family of vertical trajectories of the holomorphic quadratic differential $\varphi(z)\, dz^2$ and in each Σ_k there exists a single valued schlicht branch $\zeta = \Phi_k(z)$ of $\int \sqrt{\varphi(z)}\, dz$. Each region $D_k = \Phi_k(\Sigma_k)$ is vertically convex.

In [S] it is merely assumed that φ is analytic on Δ, instead of $\partial\Delta$, so that countably many, instead of merely finitely many Σ_k, can occur. Actually, in our use of the strip domains, the advantage of limiting ourselves to finitely many is purely didactic.

For the local and global behavior of the trajectories of holomorphic quadratic differentials we refer reader to Strebel's book ([S2]).

The following fact is important in the proof of Theorem 1.

D2. The vertical trajectories of holomorphic quadratic differential are globally geodesics in Teichmüller's metric $ds^2 = |\varphi(z)|\,|dz|^2$. Note that $\zeta = \Phi_k(z)$ is a single-valued branch of $\int\sqrt{\varphi}\,dz$ in Σ_k and that $D_k = \Phi_k(\Sigma_k)$ and $\Gamma_x = \Gamma_x^k$ is the interval which is intersection of D_k by straight line $\operatorname{Re} z = x$. Let $\theta_x = \Phi_k^{-1}(\Gamma_x)$ and $G_k = f(\Sigma_k)$. Thus θ_x is trajectory of holomorphic quadratic differential $\varphi(z)\,dz^2$. Let $\gamma_x = f(\theta_x)$. Since θ_x is a global geodesic in Teichmüller metric, we have

$$ds^2 = |\varphi(z)|\,|dz|^2,$$

$$\operatorname{length}(\Gamma_x) = \int\limits_{\theta_x} |\varphi(z)|^{1/2}|dz| \leqslant \int_{\gamma_x} |\varphi(w)|^{1/2}\,|dw|.$$

Thus we can apply Lemma 1 to the function $F_k = f\circ\Phi_k^{-1}$ and $\rho(w) = |\varphi(w)|^{1/2}$. Hence, by Lemma 1,

$$(D1) \qquad \iint\limits_{\Sigma_k} |\varphi(z)|\,dx\,dy = \operatorname{area}(D_k) \leqslant A_k B_k,$$

where

$$A_k = \left[\iint\limits_{G_k} |\varphi(w)|\,du\,dv\right]^{1/2} \quad\text{and}\quad B_k = \left[\iint\limits_{D_k} T_\nu\,d\xi\,d\eta\right]^{1/2}, \quad \nu = \nu_k = \operatorname{Belt}(F_k).$$

Using the change of variables $z = \Phi_k^{-1}(\zeta)$, we get

$${B_k}^2 = \iint\limits_{\Sigma_k} |\varphi(z)|\;T_\mu\varphi(z)\,dx\,dy.$$

Further application of the Cauchy-Schwarz lemma and (D1) gives

$$\sum_{k=1}^{n}(\varphi - \operatorname{area}(\Sigma_k)) \leqslant \sum_{k=1}^{n} A_k B_k \leqslant A\cdot B,$$

where

$$A = \left(\sum_{k=1}^{n} A_k^2\right)^{1/2} \quad\text{and}\quad B = \left(\sum_{k=1}^{n} B_k^2\right)^{1/2}$$

Now, Theorem 1 follows from the fact that

$$A = \left[\iint\limits_{\Delta} |\varphi(z)|\,dx\,dy\right]^{1/2} \quad\text{and}\quad B = \left[\int\limits_{\Delta} |\varphi| T_\mu\varphi\,dx\,dy\right]^{1/2},$$

where $\mu = \operatorname{Belt}[f]$.

IV. Extremal QC

A. Introduction. This subsections an expanded version of the lecture given by the author at The VIII Romanian-Finish Seminar, Iassy, August '99 (see [M4]). Recently, in [MM1], [BMM] and [BLMM], characterizations of unique extremality and example of unique extremal dilatation of nonconstant modulus have been obtained. Our primary purpose is to give a short exposition of some of the main result of the authors' joint papers, mentioned above, and sketch further progress in the study of a more general concept. In order to simplify exposition it is convenient to restrict mainly considerations in this review to the disc Δ.

In order to introduce and discuss a more general concept of unique extremality we need a few definitions. Let QC denote the space of all quasiconformal mappings from Δ onto itself. Two elements $f, g \in$QC are equivalent if $f = g$ on $\partial\Delta$. For a given $f \in$QC we denote the equivalence class of f by $Q_f = [f]$ or $[\mu]$, where $\mu = \mu_f$. Recall some definitions from the introduction. We also use the notation,

$$k_0([f]) = \inf\{\|\mu_g\|_\infty : g \in Q_f\}.$$

We let $L^\infty = L^\infty(\Delta)$ be the space of essentially bounded complex-valued measurable functions on Δ, and let M be the open unit ball in L^∞. For any μ in M there exists a quasiconformal solution $f : \Delta \mapsto \Delta$ of the Beltrami equation

$$(A1) \qquad \overline{\partial} f = \mu \partial f$$

unique up to a postcomposition by a Möbius transformation.

We let f^μ be the solution f of (A1) normalized by $f(i) = i$, $f(1) = 1$ and $f(-1) = -1$.

We say two elements μ_0 and μ_1 in M are equivalent if f^{μ_0} and f^{μ_1} coincide on $\partial\Delta$ and write $\mu_0 \sim \mu_1$.

The universal Teichmüller space $T = T(\Delta)$ is the space of equivalence classes of Beltrami coefficient μ in the unit ball M of the space $L^\infty = L^\infty(\Delta)$ of all essentially bounded functions on Δ. The equivalence class of the zero dilatation is the base point in T. For dilatation μ the extremal set $E = E(\mu)$ is the set where $|\mu(z)| = \|\mu\|_\infty$. We say that χ is uniquely extremal on its extremal set E if the hypothesis that μ is equivalent to χ, in its Teichmüller class together with the condition $\|\mu\|_\infty \leqslant \|\chi\|_\infty$, imply that $\mu = \chi$ a.e. on E. One verifies easily that if χ is uniquely extremal on its extremal set E and if the measure of E is positive, then χ is an extremal dilatation.

In order to prove that $f \in$QC is extremal (or uniquely extremal) we need estimates which allow us to compare Beltrami dilatation μ_f of f, with μ_g, for the other $g \in Q_f$. It appears that the main inequality of Reich and Strebel is a major tool in theory of extremal quasiconformal mappings. Using the main inequality in [BLMM] we have derived an inequality, which we called Delta inequality, and we have shown that the Delta inequality is suitable for studying unique extremality.

In order to give a characterization of a given dilatation χ, which is uniquely extremal on its extremal set $E = E(\chi)$, we make a variation χ_r of χ on a compact set $K \subset E$. It turns out that $[\chi_r]$ is a Strebel point. Using Strebel Frame Mapping Criterion and the main inequality, we show that for $r \in (0, r_0)$, where r_0 is a positive number small enough, there is a unit vector $\varphi = \varphi_r$ such that

$$\delta(\varphi) = \delta_\chi(\varphi) = \|\chi\|_\infty - \int_\Delta \chi\varphi \, dx\, dy \leqslant 2r \int_K |\varphi|.$$

Here Q is the subspace of $L^1 = L^1(\Delta)$ consisting of holomorphic function in Δ.

Now, if $\lambda = \lambda_\chi$ is the linear functional defined by $\lambda(\varphi) = \iint_\Delta \chi\varphi\, dx\, dy$, we can elementary show that

(1) λ_χ has a unique norm-preserving extension from Q to $Q_{\tilde\chi}$

Here $\tilde\chi$ is defined by $\tilde\chi = \overline{\chi}$ on E, $\tilde\chi = 0$ on $\Delta \setminus E$ and $Q_{\tilde\chi}$ is the smallest subspace of L^1 which contains $Q \cup \{\tilde\chi\}$.

Proof that (1) implies that χ is uniquely extremal on E is based on the Delta inequality. Analysis of the proof of Hahn-Banach Theorem shows that (1) is equivalent to

(2) There exists sequence $\{u_n\}$ in Q such that $\lambda(u_n) = \lambda(\tilde\chi) + \|\chi\|_\infty \|\tilde\chi - u_n\| + o(1)$, where $o(1) \to 0$, when $n \to \infty$.

Hence, using the classical result that L^1-convergence of a sequence of functions $\{f_n\}$ implies that there exists subsequence $\{f_{n_k}\}$, which converges a.e. on the corresponding set, and the Delta inequality, we can prove the following result.

Theorem A. *χ is uniquely extremal on its extremal set E if and only if*
emph(3) χ satisfies Reich condition on E,
that is, there exists sequence $\{\varphi_n\}$ in Q, such that
(a) $\delta(\varphi_n) = \|\varphi_n\|\|\chi\|_\infty - \operatorname{Re}\int_\Delta \varphi_n\chi \to 0$
(b) $\lim_{n\to\infty} \inf |\varphi_n(z)| > 0$ *for almost all z in $E(\chi)$.*

This criterion has many applications concerning precise characterizations (Theorem G2 and Corollary G3) and removable properties (Theorem G3 and Theorem G4) of unique extremal dilatation of Teichmüller type. Also, this criterion can be used to construct unique extremal dilatation of nonconstant modulus.

B. Extremal dilatation. In this section we give a short report concerning extremal mappings. The interested reader can learn more about extremal mappings from Strebel's survey article [S6], Reich's papers [Re8], [Re9] and Earle-Li Zhong's [ELi] , all of which we highly recommend. Extremal mappings have been one of the main topics in the theory of quasiconformal mappings, since its earliest days, when Grötzsch solved the extremal problem for two rectangles. In order to discuss them we need to review some familiar definitions.

A homeomorphism f from a domain G onto another is called quasiconformal if f is ACL (absolutely continuous on lines) in G and $|f_{\bar z}| \leqslant k|f_z|$ a.e. in G, for some real number k, with $0 \leqslant k < 1$.

It is well known that if f is a quasiconformal mapping defined on the region G, then the function f_z is nonzero a.e. in G. The function $\mu_f = f_{\bar z}/f_z$ is therefore a well defined bounded measurable function on G, called the complex dilatation or Beltrami coefficient of f.

The positive number

$$K(f) = \frac{1 + \|\mu_f\|_\infty}{1 - \|\mu_f\|_\infty}$$

is called the maximal dilatation of f. We say that f is K-quasiconformal if f is a quasiconformal mapping and $K(f) \leqslant K$. Let QC denote the space of all quasiconformal mappings from Δ onto itself. Two elements $f, g \in$QC are equivalent if $f = g$ on $\partial\Delta$. For a given $f \in$QC we denote the equivalence class of $f \in$QC by $Q_f = [f]$ or $[\mu]$, where $\mu = \mu_f$. Also we use notation

$$k_0([f]) = \inf\{\|\mu_g\|_\infty : g \in Q_f\} \quad \text{and} \quad K_0([f]) = \frac{1 + k_0}{1 - k_0}.$$

We let $L^\infty = L^\infty(\Delta)$ be the space of essentially bounded complex-valued measurable functions on Δ, and let M be the open unit ball in L^∞. For any μ in M there exists a quasiconformal solution $f : \Delta \mapsto \Delta$ of the Beltrami equation

$$\overline{\partial} f = \mu \, \partial f \tag{1}$$

unique up to a postcomposition by a Möbius transformation.

We let f^μ be the solution f of (1) normalized by $f(i) = i$, $f(1) = 1$ and $f(-1) = -1$. Two elements μ_0 and μ_1 in M are equivalent if f^{μ_0} and f^{μ_1} coincide on $\partial\Delta$. For given $\mu \in M$ the equivalence class $[\mu]$ contains at least one element μ_0 such that $\|\mu_0\|_\infty = \inf\{\|\nu\|_\infty : \nu \in [\mu]\}$. Such a μ_0 is referred to as an extremal complex dilatation and $f_0 = f^{\mu_0}$ as an extremal quasiconformal mapping (abbreviated EQC mapping).

As we mentioned, we restrict mainly considerations in this review to the disc and only consider a few examples concerning the other Jordan domains.

For discussion concerning subregions of the plane, which are not necessarily simply-connected, we refer the interested reader to the paper [ELi], which we follow in this section.

Let z_i, $1 \leqslant i \leqslant 4$, be four distinct points on the unit circle S^1, and let w_i, $1 \leqslant i \leqslant 4$, be their images under some sense-preserving homeomorphism h of the closed unit disc Δ onto itself. Let S be the set of all quasiconformal mappings g of the open unit disc Δ onto itself which maps z_i to w_i, $1 \leqslant i \leqslant 4$.

Let ϕ and ψ be conformal mappings of the two discs onto horizontal rectangles, R and R', which map the distinguished points (vertices) on vertices of rectangles. Let A_K be the horizontal stretching of R onto R' defined by $A_K(\zeta) = K\xi + i\eta$, where $\zeta = \xi + i\eta$. Then $f = \psi^{-1} \circ A_K \circ \phi$ is the unique extremal mapping.

It is easy to compute the Beltrami dilatation of f, that we have just described. One finds that

$$\mu_f = k\frac{\overline{\phi'}}{\phi'} = k\frac{|\phi'|^2}{(\phi')^2}. \tag{2}$$

Note that the holomorphic function $(\phi')^2$ belongs to $Q(\Delta)$ since its L^1 norm equals the area of rectangle R.

Teichmüller discovered that many extremal quasiconformal mappings have Beltrami dilatation whose form resembles (2).

In recognition of the importance of that discovery, a Beltrami coefficient μ in a plane region G is called a Teichmüller dilatation if there are number $k \in [0,1)$ and a holomorphic function $\varphi \in Q(G)$, not identically zero, such that $\mu = k|\varphi|/\varphi$, a.e. in G.

If we do not require that φ is integrable we will say that μ is Teichmüller dilatation in general sense (or μ has Teichmüller type). Here, for a given domain $G \subset \mathbb{C}$, by $Q(G)$ we denote the space of all holomorphic function φ in $L^1(G)$. A quasiconformal mapping f whose Beltrami coefficient is a Teichmüller dilatation is called a Teichmüller mapping.

Let Q be the Banach space consisting of holomorphic functions φ, belonging to $L^1 = L^1(\Delta)$, with norm

$$\|\varphi\| = \iint_\Delta |\varphi(z)|\, dx\, dy \, < \infty, \; \varphi \in Q.$$

For $\mu \in L^\infty$ we consider the linear functional $\Lambda_\mu(\varphi) = (\mu, \varphi)$, $\varphi \in Q$, where

$$(\mu, \varphi) = \iint_\Delta \mu(z)\varphi(z)\, dx\, dy,$$

and denote by $\|\mu\|_* = \|\Lambda_\mu\|$ the norm of μ as an element of the dual space of Q. For $\mu \in L^\infty$ we say that it satisfies the Hamilton-Krushkal condition if $\|\mu\|_* = \|\mu\|_\infty$.

We are now ready to state the main result about extremal complex dilatations.

Theorem HKRS. (Hamilton–Krushkal and Reich–Strebel) *Let $\mu \in M$. A necessary and sufficient condition that f^μ is an EQC mapping is that*

$$\|\mu\|_* = \|\mu\|_\infty. \tag{3}$$

We are going to prove the necessity of Hamilton-Krushkal condition for a quasiconformal mapping to be extremal in its Teichmüller class (see Theorem HK bellow). For the proof we need two lemmas.

Let N denote the subspace of $L^\infty(\Delta)$ which is orthogonal to Q. Differentials which belong to N are called infinitesimally trivial.

For the proof of the next Lemma see Lehto [L2].

Lemma B1. *Let $\nu \in N$ and $\|\nu\|_\infty < 2$. Then, for $0 \leqslant t \leqslant 1/4$, there is a $\sigma_t \in [t\nu]$ such that $\|\sigma_t\|_\infty \leqslant 12\, t^2$.*

Lemma KRS. *Let f^μ be extremal. If μ and χ represent the same infinitesimal equivalence class, then $\|\mu\|_\infty \leqslant \|\chi\|_\infty$.*

Proof. Set $k_0 = \|\mu\|_\infty$, $k = \|\chi\|_\infty$. We assume that $k < k_0$ and prove that f^μ cannot be extremal. Writing $\nu = \mu - \chi$, we first prove that for $f^\lambda = f \circ (f^{t\nu})^{-1}$ has a smaller maximal dilatation than f^μ. Since

$$\lambda(\zeta) = \frac{\mu - t\nu}{1 - t\mu\overline{\nu}} \frac{1}{\theta}, \quad \text{where } \theta = \overline{p}/p, \; p = \partial f^{t\nu},$$

direct calculation gives $|\lambda(\zeta)| = |\mu(z)| - At + O(t^2)$, where

$$\zeta = f^{t\nu}(z), \; A = A_{\mu,\chi}(z) = \frac{1 - |\mu(z)|^2}{|\mu(z)|} \operatorname{Re} \mu(z)\overline{\nu(z)}$$

and the remainder term $O(t^2)$ is uniformly bounded in z.

Write $E_1 = \{z \in \Delta : |\mu(z)| < (k + k_0)/2\}$, $E_2 = \Delta \setminus E_1$. In E_2,

$$\operatorname{Re} \frac{\mu \overline{\nu}}{|\mu|} = |\mu| - \operatorname{Re} \frac{\mu}{|\mu|} \overline{\nu} \geqslant \frac{k + k_0}{2} - k = \frac{k_0 - k}{2}$$

and therefore

$$A \geqslant \frac{1}{2}(1 - k_0^2)(k_0 - k).$$

Hence, $|\lambda(\zeta)| < k_0 - m_0 t$. But, if σ_t is as in Lemma B1, then $f^\tau = f^\lambda \circ f^{\sigma_t}$ and f^μ belong to the same Teichmüller class. We have

$$\|\tau\|_\infty \leqslant \frac{\|\lambda\|_\infty + \|\sigma_t\|_\infty}{1 - \|\sigma_t\|_\infty} < \frac{k_0 - m_0 t + 12t^2}{1 - 12t^2}.$$

Theorem HK. *If f^μ is extremal in its equivalence class, then*

$$\|\mu\|_* = \|\mu\|_\infty. \tag{4}$$

Proof. By the Hahn-Banach theorem and Riesz representation theorem there exists $\chi \in M$ such that μ and χ represent the same infinitesimal class and $\|\chi\|_\infty = \|\Lambda_\mu\| = \|\mu\|_*$. Since $\|\chi\|_\infty = \|\mu\|_* \leqslant \|\mu\|_\infty$, the equality (4) follows from Lemma KRS.

The necessary condition (4) for μ to be extremal is also sufficient. Reich and Strebel proved this using their "Main Inequality". We find it is convenient to formulate Hamilton-Krushkal's condition in terms of Hamilton sequences.

Definition B1. Let μ_f be the Beltrami coefficient of some quasiconformal mapping f of the unit disc Δ onto itself. A Hamilton sequence for μ_f, is a sequence in Q, such that $\|\varphi_n\| = 1$, for all n, and $\lim_{n \to \infty} (\mu, \varphi_n) = \|\mu\|_\infty$.

Now we can state theorem of Hamilton-Krushkal and Reich–Strebel in the form

Theorem B1. *Let f be a quasiconformal mapping of the unit disc Δ onto itself, and let μ_f be its Beltrami coefficient. Then f is extremal in its class $[f]$ if and only if μ_f has a Hamilton sequence.*

Corollary B1. *Every Teichmüller mapping is extremal in its equivalence class.*

Proof. If f is a Teichmüller mapping, then its Beltrami coefficient μ_f can be written in the form $\mu_f = k|\varphi|/\varphi$, with $0 < k < 1$, $\varphi \in Q$ and $\|\varphi\| = 1$. The sequence $\{\varphi_n\}$, with $\varphi_n = \varphi$, for all n, is obviously a Hamilton sequence for μ_f.

Not all extremal quasiconformal mappings are Teichmüller mappings. The first counterexample occurs in the famous paper [BA].

Example B1. Let H be the upper half plane and $K > 1$. In the section 5 of [BA] it is shown that the quasiconformal mapping $f(z) = z\,|z|^{K-1}$ of H onto itself is extremal in its class. A simple calculation shows that

$$\mu_f(z) = k \frac{|\psi(z)|}{\psi(z)}, \quad z \in H,$$

where $k = (K-1)/(K+1)$ and $\psi(z) = z^{-2}$. This has the same form as Teichmüller mapping, but ψ is not Teichmüller mapping, because ψ is not an integrable function on the upper half plane.

The question whether f is uniquely extremal mapping in its class was not considered in [BA]. One method for studying extremal mappings that are not Teichmüller mappings is to use degenerating Hamilton sequences.

Let μ_f be the Beltrami coefficient of the quasiconformal mapping f. The Hamilton sequence $\{\varphi_n\}$ for μ_φ is degenerating if φ_n converges zero uniformly on compact subsets on Δ as $n \to \infty$. The connection between degenerating Hamilton sequences and Teichmüller mappings is given by the following.

Lemma B2. (Reich and Strebel) *If μ_f has a Hamilton sequence that does not degenerate, then f is a Teichmüller mapping.*

For a proof of this Lemma see for example Earle–Li Zhong [EL] (see also Lehto [L2]). Lemma B2 shows that it is desirable to find geometric condition on an extremal mapping that will prevent its Beltrami coefficient from having a degenerating Hamilton sequences. Strebel's Frame Mapping Criterion provides such conditions in terms of the boundary dilatation, which we shall now define.

The boundary dilatation $H([\mu])$ of the Teichmüller class of μ is the infimum over all elements ν in the equivalence class of μ in T of the quantity

$$\frac{1+h^*(\nu)}{1-h^*(\nu)}.$$

Here $h^*(\nu)$ is the infimum over all compact subsets K contained in Δ of the essential supremum of Beltrami coefficient $\nu(z)$ as z varies over $\Delta \setminus K$. As usual, we let

$$H^*(\mu) = \frac{1+h^*(\mu)}{1-h^*(\mu)}.$$

For $f \in$QC, also, we define $H^*(f) = H^*(\mu_f)$ and $H([f]) = H([\mu_f])$.

Theorem B2. (Strebel Frame Mapping Criterion) *Let $f \in QC$ and let f be extremal in its class Q_f. If $H([f]) < K(f)$, then*

(a) μ_f has no degenerating Hamilton sequences

(b) f is a Teichmüller mapping.

For a proof of this theorem see for example Gardiner [Ga].

Example B2. (Strebel's chimney) In [S1], Strebel made another breakthrough by constructing the first example of a nonuniquely extremal Beltrami coefficient. Strebel considered the plane region

$$V = \{z = x + iy : |x| < 1\} \cup \{z = x + iy : y < 0\},$$

now known as Strebel's chimney. For every real number $K > 1$, the quasiconformal homeomorphism $f_K(z) = x + iKy$ of V is extremal in its class $\tau_K \in T(V)$. On the other hand, τ_K contains infinitely many distinct extremal mappings. For instance, $h_L(z) = f_K(z)$, for $y \geqslant 0$, and $h_L(z) = f_L(z)$, for $y < 0$, is extremal in τ_K for every $L \in [1/K, K]$.

C. Unique extremality. In this section we shortly discuss results of authors' joint paper [BLMM] (see also [MM1] and [BMM]). For studying unique extremality it is convenient to use the following result.

Proposition C1. *Let φ be conformal mapping of the domain D onto V. Then ν is uniquely extremal on V if and only if $\mu(z) = \nu(\varphi(z))\overline{\varphi'(z)}/\varphi'(z)$ is uniquely extremal on D.*

Strebel has proved in [S3] that the horizontal stretching $A(w) = A_K(w) = Ku + iv$, $K > 1$, in $V = \{w : |v| < \pi/4\}$ is the unique extremal in its equivalence class. The proof of Strebel's result has also been given in [ELi], using Reich's method (see also [Re9]).

Using conformal mapping $w = \varphi(z) = \frac{1}{2}\ln z - \frac{\pi}{4}$ of H onto V one can show that the Beurling-Ahlfors mapping f is the uniquely extremal mapping in its class.

Ahlfors and Bers showed that T has a complex structure with tangent space at the base point isomorphic to Banach space Q^*. Two tangent vectors μ and ν in the tangent space to M determine the same tangent vector in T if and only if

$$\int_\Delta \varphi\mu = \int_\Delta \varphi\nu, \text{ for all } \varphi \in Q.$$

If μ and ν have this property, we say that they represent the same Teichmüller infinitesimal equivale e class or, more briefly, that they are infinitesimally equivalent. The space of equivalence classes is denoted by B. A given μ is said to be extremal in its infinitesimal Teichmüller class if $\|\mu\|_\infty \leqslant \|\nu\|_\infty$, for any ν infinitesimally equivalent to μ.

Recall that Hamilton, Krushkal, Reich and Strebel showed that a Beltrami coefficient ν in M is extremal in its class in T if and only if ν is extremal in its class in B. It was natural to consider whether the analogous statement holds for the unique extremality. In several articles Reich showed that in many special situations the two notions of unique extremality coincide and conjectured that the notions may coincide in general. In [BLMM] (see also [MM1] and [BMM]) we have recently proved the answer to this conjecture is affirmative.

Theorem C1. (The Equivalence Theorem I) *μ is uniquely extremal in its Teichmüller class if and only if μ is uniquely extremal in its infinitesimally class.*

Proof of this theorem is based on estimates which allow us to compare two Beltrami coefficients μ and ν in the same global equivalence class and two Beltrami differentials in the same infinitesimal equivalence class. These estimates generalize Reich's Delta inequality for Beltrami differentials in the same equivalence class (see [R8]). Unlike Reich's forms of the Delta inequalities, our forms do not require either one of the Beltrami coefficients to have constant absolute value.

The generalized Delta inequality is our first step towards obtaining the criterion for the unique extremality of Beltrami differentials. The next important step is the analysis of the proof of Hahn-Banach theorem and its applications to our setting. In particular, we obtain the following necessary and sufficient criterion for the unique extremality of given Beltrami coefficient χ.

Theorem C2. (Characterization Theorem I) *Beltrami coefficient χ is uniquely extremal if and only if for every admissible variation η of χ there exist a sequence φ_n in $A(R)$ such that*

(a) $\delta(\varphi_n) = \|\varphi_n\|\|\eta\|_\infty - \operatorname{Re}\int_R \varphi_n\eta \to 0$

(b) $\lim\limits_{n\to\infty} \inf |\varphi_n(z)| > 0$, *for almost all z in $E(\eta)$.*

Here, an admissible variation η of χ is any Beltrami differential that does not increase the L^∞-norm of χ, and which is allowed to differ from χ only on the set where $|\chi(z)| \leqslant s < \|\chi\|_\infty$, where s is a constant, and the extremal set $E(\eta)$ is the set where $\eta(z) = \|\eta\|_\infty$. This criterion is analogous to the Hamilton-Krushkal, Reich–Strebel necessary and sufficient criterion for the extremality. Namely, χ is extremal if and only if there is a sequence φ_n of holomorphic quadratic differential of norm 1 such that

$$\|\chi\|_\infty - \operatorname{Re}\int_R \eta\varphi_n \to 0.$$

This criterion is among listed in the theorem in Section 11, in [BLMM], which we called the Characterization Theorem. The Characterization Theorem applies to many interesting situations. For instance, we can say precisely when a Beltrami differential of the form $k|\varphi(z)|/\varphi(z)$, with φ a holomorphic quadratic differential with $\|\varphi\| = \infty$, is uniquely extremal.

There are many examples of extremal Beltrami differentials with nonconstant modulus, but all examples of uniquely extremal Beltrami differentials known up to our papers [BLMM] and [BMM] were of the general Teichmüller type. Moreover, many results obtained studying the extremal problems speak in favour of the conjecture that all uniquely extremal Beltrami differentials μ satisfy $|\mu(z)| = \|\mu\|_\infty$, for almost all z. Surprisingly, we disprove this conjecture and show that there are uniquely extremal Beltrami differentials with nonconstant modulus.

D. The main inequalities. Let Δ denote the unit disc,

$$S_\mu\varphi = \left|1 - \mu(z)\frac{\varphi(z)}{|\varphi(z)|}\right|^2 \quad \text{and} \quad T_\mu\varphi(z) = \frac{S_\mu\varphi}{1-|\mu(z)|^2} = \frac{\left|1 - \mu(z)\frac{\varphi(z)}{|\varphi(z)|}\right|^2}{1-|\mu(z)|^2}.$$

We will refer to the following result as the Reich–Strebel inequality or the Main Inequality.

Theorem RS. (Reich and Strebel). *Suppose that f is a quasiconformal homeomorphism of Δ onto itself which is the identity on $\partial\Delta$. Then, with $\mu = \mu_f$*

$$\iint\limits_\Delta |\varphi(z)|\,dx\,dy \leqslant \iint\limits_\Delta |\varphi(z)|T_\mu\varphi(z)\,dx\,dy,$$

for every analytic integrable function φ on Δ.

Various forms of this result play a major role in the theory of quasiconformal mappings and have many applications.

For applications to extremal and uniquely extremal quasiconformal mappings, we refer the interested reader to the book by Gardiner ([G]), and for some recent results to [MM1], [BMM], [BLMM], [Re3] and [Re9] .

Let f and g be two equivalent quasiconformal mappings on Δ and let

$$\mu = \mu_f = \mathrm{Belt}[f],\ \alpha = \mu_{f^{-1}} \circ f,\ \beta = \mu_{g^{-1}} \circ f \text{ and } \tau = \mu \frac{\beta}{\alpha}.$$

Then $g^{-1} \circ f$ is the identity on $\partial\Delta$, and, if we apply the Reich–Strebel inequality to $F = g^{-1} \circ f$, we get

$$1 \leqslant \iint_{\Delta} |\varphi(z)| T_\mu \varphi T_{-\tau\theta}\varphi \, dx\, dy, \tag{1}$$

where $\theta = (1 - \overline{\mu\varphi}/|\varphi|)(1 - \mu\varphi/|\varphi|)^{-1}$ and $\varphi \in Q$, $\|\varphi\| = 1$. Note that $\alpha \doteq -\mu p/\overline{p}$, $\tau = -\overline{p}/p\beta$ and

$$T_{-\tau\theta}\varphi = \frac{S_{-\tau\theta}\varphi}{1 - |\beta|^2}.$$

Now, we are going to state two consequences of the Main Inequality, known as the fundamental Reich–Strebel inequalities (inequalities (2) and (3) below). If $K_0 = K_0([\mu])$ and $g \in Q_f$ is an extremal quasiconformal mapping, then inequality (1) yields

$$\frac{1}{K_0} \leqslant \iint_{\Delta} |\varphi| T_\mu \varphi \, dx\, dy. \tag{2}$$

Suppose now that μ is a Teichmüller differential, i.e., $\mu = k_0|\varphi_0|/\varphi_0$ for some $\varphi_0 \in Q$, with $\|\varphi_0\| = 1$ and $0 < k_0 < 1$. Then $\theta = 1$ and $T_\mu\varphi_0 = K_0^{-1}$, where $K_0 = \frac{1+k_0}{1-k_0}$. Suppose that ν equivalent to $k_0|\varphi_0|/\varphi_0$, where $\|\varphi_0\| = 1$.

It means that there exists $g = f^\nu$, which is is equivalent to $f = f^\mu$. Therefore, the inequality (1) becomes

$$K_0 \leqslant \iint_{\Delta} |\varphi_0| T_{-\tau}\varphi_0 \, dx\, dy \tag{3}$$

Since $\|\tau\|_\infty = \|\beta\|_\infty = \|\nu\|_\infty$, it follows from (3) that $K_0 \leqslant \frac{1+\|\nu\|_\infty}{1-\|\nu\|_\infty}$, which implies $k_0 \leqslant \|\nu\|_\infty$ and, therefore, $k_0|\varphi_0|/\varphi_0$ has minimal norm among all equivalent Beltrami dilatations ν. Moreover, if $k_0 = \|\nu\|_\infty$, then the inequality (3) yields

$$K_0 \leqslant \iint_{\Delta} |\varphi_0| T_{-\nu}\varphi_0 \, dx\, dy \leqslant K_0, \tag{4}$$

and so (4) is an equality and this obviously implies first that $\alpha = \beta$, i.e., $f^{-1} = g^{-1}$. Hence $f = g$ and therefore $\nu = k_0|\varphi_0|/\varphi_0$ almost everywhere.

We have proved the following theorem.

Theorem T. (Teichmüller Uniqueness Theorem) *Suppose that $\mu = k|\varphi|/\varphi$, where $0 < k < 1$ and φ is an element of norm 1 in Q. Then $\mu = k|\varphi|/\varphi$ is uniquely extremal in its Teichmüller class on Δ.*

Note that Teichmüller Uniqueness Theorem also follows directly from the Delta inequality, which will be considered in the next subsection.

E. Delta inequality. One verifies easily that the main inequality of Reich and Strebel can be stated in the following form.

Theorem E1. (Reich and Strebel) textitSuppose that f is a quasiconfomal homeomorphism of Δ onto itself which is the identity on $\partial\Delta$.Then, with $\mu = \mu_f$

$$\operatorname{Re} \iint\limits_{\Delta} \frac{\mu}{1-|\mu|^2} \varphi \, dx \, dy \leqslant \iint\limits_{\Delta} \frac{|\mu|^2}{1-|\mu|^2} |\varphi| \, dx \, dy, \text{ for all } \varphi \in Q.$$

A simple calculation shows that this inequality is equivalent to the inequality (1) (see below), which is a starting point in the proof of the Delta inequality.

Theorem E2. (The Delta inequality in T) *Let μ and ν belong to the same class in T, $f = f^{\mu}$, $g = f^{\nu}$, $\alpha = \mathrm{Belt}[f^{-1}] \circ f$, $\beta = \mathrm{Belt}[g^{-1}] \circ f$, $\rho = |\alpha(z) - \beta(z)|^2$ and $I = I(\varphi) = \int_{\Delta} \rho |\varphi|$. If $\|\nu\|_{\infty} \leqslant k = \|\mu\|_{\infty}$, then $I(\varphi) \leqslant C\delta_{\mu}(\varphi)$, $\varphi \in Q$, where C is a constant which depends only on $k = \|\mu\|_{\infty}$.*

Proof. We will prove this result under additional hypothesis that $|\mu|$ is bounded from below by a positive constant s, for almost every z in Δ. For a complete proof we refer the interested reader to [BLMM]. Let $P = (\alpha - \beta)(1 - \alpha\overline{\beta})\alpha^{-1}$ and $Q = (1 - |\alpha|^2)(1 - |\beta|^2)$. Then for any $\varphi \in Q$ we have as an easy consequence of the Main Inequality

$$\operatorname{Re} \int_{\Delta} \frac{P}{Q} \mu\varphi \leqslant \int_{\Delta} \frac{\rho}{Q} |\varphi|. \tag{1}$$

In order to get an estimate involving $\delta\{\varphi\}$ add to both sides $l = \int_{\Delta} |\alpha| \frac{\operatorname{Re} P}{Q} |\varphi|$. We get

$$\int_{\Delta} \frac{|\alpha| \operatorname{Re} P - \rho}{Q} |\varphi| \leqslant \operatorname{Re} \int_{\Delta} \frac{P}{Q} (|\alpha||\varphi| - \mu\varphi).$$

Furthermore, $|\alpha| \operatorname{Re} P - \rho = A\rho + B(|\alpha| - |\beta|)$, where $A = 2^{-1}|\alpha|^{-1}(1 - |\alpha|)^2$ and $B = 2^{-1}|\alpha|^{-1}(|\alpha| + |\beta|)(1 - |\alpha|^2)$. Hence, the inequality (1) can be rewritten in the form

$$\int_{\Delta} \frac{A\rho}{Q} |\varphi| \leqslant J(\varphi) + s(\varphi), \tag{2}$$

where

$$J(\varphi) = \operatorname{Re} \int_{\Delta} \frac{P}{Q} (|\mu||\varphi| - \mu\varphi) \quad \text{and} \quad s(\varphi) = \int_{\Delta} \frac{B}{Q} (|\beta| - |\alpha|) |\varphi|.$$

Since μ is bounded from below by a positive constant $s > 0$ in Δ, A and B are bounded, using (2) and the estimate $|\beta| - |\alpha| \leqslant k - |\mu|$, one can show that there is a constant c, which depends only on $k = \|\mu\|_{\infty}$ and s, such that $I \leqslant c(I_1 + \tau\{\varphi\})$, where

$$I_1 = \int_{\Delta} \sqrt{\rho} ||\mu\varphi| - \mu\varphi| \quad \text{and} \quad \tau = \tau\{\varphi\} = \int_{\Delta} (k - |\mu|) |\varphi|.$$

Using Cauchy-Schwarz inequality and the identity $||w|-w|^2 = 2|w|(|w|-\operatorname{Re} w)$, we obtain $I_1 \leqslant cI^{1/2}\delta^{1/2}$. Using this and the inequality $\tau \leqslant \delta$, we get $I \leqslant c(I^{1/2}\delta^{1/2} + \delta)$, and therefore $I \leqslant C(k,s)\delta$.

F. More general concept of unique extremality. We say that $\chi \in M$ is uniquely extremal on its extremal set E if $|E| > 0$ and the hypothesis that μ is equivalent to χ, in its Teichmüller class together with the condition $\|\mu\|_\infty \leqslant \|\chi\|_\infty$, imply that $\mu = \chi$ a.e. on E.

We say that $\chi \in L^\infty$ satisfies unique extension property on its extremal set E (or we say that χ is unique extremal on E in its infinitesimal Teichmüller class B), if $|E| > 0$ and the hypothesis that μ is equivalent to χ, in its infinitesimal Teichmüller class B together with the condition $\|\mu\|_\infty \leqslant \|\chi\|_\infty$, imply that $\mu = \chi$ a.e. on E.

Theorem F1. (The Equivalence Theorem II) *Let $\chi \in M$ and E be its extremal set. Then the following conditions are equivalent*

(a) *χ is uniquely extremal on E*

(b) *χ satisfies unique extension property on E.*

Note that an immediate consequence of this result is the Equivalence Theorem when dilatation has constant absolute value. First, we give a few definitions and lemmas, on which the proof is based.

Now, we will introduce the variation property and prove Lemma VT, Lemma V and two other lemmas (we develop them in T and B) which utilize variational property.

Definition F1. Let $\mu \in L^\infty$, $k = \|\mu\|_\infty < 1$ and E be its extremal set. We say that μ satisfies variational property on E in T if for each compact subset $K \subset E$ and for each $r > 0$

$$(1+r)k_0([\mu_r]) > k_0([\mu]),$$

where $\mu_r = \dfrac{\mu}{1+r}$ in K^c and $\mu_r = \mu$ in K.

In a parallel manner, we define the variational property in B.

Definition F2. Let $\mu \in L^\infty$, $k = \|\mu\|_\infty$ and E be its extremal set. We say that μ satisfies variational property on E in B if for each compact subset $K \subset E$ and for each $r > 0$

$$\|\mu_r\|_* > \|\mu\|_*,$$

where $\mu_r = \mu$ in K^c and $\mu_r = (1+r)\mu$ in K.

Lemma VT1. *If μ is uniquely extremal on its extremal set E, then μ satisfies variational property on E in T.*

Proof. Let $K \subset E$ be a set of positive measure and let μ_r be the variation of μ to K. Assume that $H(\mu_r) = K_0(\mu_r)$. Then, there exists ν in Teichmüller class of μ_r such that $\|\nu\|_\infty \leqslant s_0$.

Let $g = f^\nu \circ g_r$, where $g_r = (f^{\mu_r})^{-1} \circ f^\mu$. As in [BLMM] the reader can verify that $K(g) \leqslant K(f^\mu)$. Using that g_r converges uniformly to the identity on K, when $r \to 0+$, one can show that the set $A = K \cap g_r(K)$ has positive measure if $r \in (0, r_0)$

for some positive r_0. This means that f and g are distinct on the set $B = g_r^{-1}(A)$. Since $B \subset E$ has positive measure and f is uniquely extremal on E this yields a contradiction.

Lemma VT2. *Let μ satisfies variational property on E in T and let $K \subset E$ be compact set of positive measure on which $|\mu| = \|\mu\|_\infty < 1$. Then for each $r > 0$ there is a unit vector $\varphi \in Q$ such that*

$$\delta_\mu(\varphi) \leqslant 2r \int_K |\varphi|. \tag{1}$$

Proof. Let $k_0 = k_0([\mu])$, $s_0 = k_0/(1+r)$, $\mu_r = \frac{\mu}{1+r}$ in K^c and $\mu_r = \mu$ in K. Since $H([\mu_r]) \leqslant H^*(\mu_r) \leqslant \frac{1+s_0}{1-s_0}$ and by Lemma VT, $K_0([\mu_r]) > \frac{1+s_0}{1-s_0}$, we conclude that the extremal dilatation K_0 is strictly greater then the boundary dilatation H. Thus $[\mu_r]$ is a Strebel point in T and by Strebel's frame mapping theorem there exists $s_r = k_0([\mu_r]) > k_0$ and a unit vector $\varphi \in Q$ such that μ_r and $s_r\frac{|\varphi|}{\varphi}$ are equivalent in T.

Therefore, by Reich–Strebel's second fundamental inequality (see the inequality (3) in Section C and also [Ga]),

$$\frac{1+s_0}{1-s_0} \leqslant \frac{1+s_r}{1-s_r} \leqslant \int_\Delta |\varphi| \frac{|1+\mu_r\frac{\varphi}{|\varphi|}|^2}{1-|\mu_r|^2}.$$

A simple calculation as in [BLMM] gives (1).

Now, we are going to prove that condition (a) implies the condition (b) in Theorem F1.

Proof. Let μ be uniquely extremal on its extremal set E in its Teichmüller class and $k_0 = k_0([\mu])$. Suppose that μ does not satisfy the condition (b). Hence, there exists ν, distinct from μ on E, such that μ and ν belong to the same class in B and $\|\nu\|_\infty \leqslant k_0$. Therefore, there exist $\epsilon \in (0, k_0)$ and a compact set $K \subset E$ of positive measure such that $|\mu(z) - \nu(z)| \geqslant 2\epsilon$, a.e. on K. Since

$$\delta_\mu(\varphi) = k_0\|\varphi\| - \operatorname{Re}\int_\Delta \frac{\mu+\nu}{2}\varphi,\ \varphi \in Q \quad \text{and} \quad \left|\frac{\mu+\nu}{2}\right| \leqslant d, \text{ a.e. on } K,$$

where $d = \sqrt{k_0^2 - \epsilon^2}$, we conclude that $(k_0 - d)\|\varphi\|_K \leqslant \delta_\mu(\varphi)$. Here we use the notation $\|\varphi\|_K = \iint_K |\varphi|\, dx\, dy$. Hence, using Lemma VT1 and Lemma VT2 one can get a contradiction.

Lemma V1. *If μ satisfies unique extension property on its extremal set E, then μ satisfies variational property on E in B.*

Proof. Contrary, suppose that $\|\mu_r\|_* \leqslant k = \|\mu\|_\infty$ for some $r > 0$ and some compact set $K \subset E$, where $\mu_r = \mu$ in K^c and $\mu_r = (1+r)\mu$ in K. Then there is a non-zero annihilator $\eta \in N$ such that $\|\mu_r + \eta\|_\infty \leqslant k$.

Let $\mu_1 = \mu + \epsilon\eta$, where $\epsilon = (1+r)^{-1}$. By using similarity of the triangles, one can check that $\|\mu_1\| \leqslant k$. Since $\mu_1 \in [\mu]$, we conclude that $\mu_1 = \mu$, i.e., $\eta = 0$. Thus we have a contradiction.

Lemma V2. *Let μ satisfy the variational property on E in B and let $K \subset E$ be a compact set of positive measure on which $|\mu| = \|\mu\|_\infty = k$. Then for each $r > 0$ there is a unit vector $\varphi \in Q$ such that*

$$\delta_\mu(\varphi) \leqslant kr \int_K |\varphi|.$$

Proof. Let $\mu_r = \mu$ in K^c, $\mu_r = (1+r)\mu$ in K and let λ_r be a linear functional defined by $\lambda_r(\varphi) = \operatorname{Re}(\mu_r, \varphi)$, $\varphi \in Q$. Since μ satisfies variational property in B, there exists a unit vector $\varphi \in Q$ such that $\lambda_r(\varphi) \geqslant k$. Therefore,

$$\delta_\mu(\varphi) = k - \lambda(\varphi) \leqslant \lambda_r(\varphi) - \lambda(\varphi).$$

Since

$$\lambda_r(\varphi) - \lambda(\varphi) = \operatorname{Re}(\mu_r - \mu, \varphi) = \operatorname{Re} \int_K r\mu,$$

one can find that

$$\delta_\mu(\varphi) \leqslant |\lambda_r(\varphi) - \lambda(\varphi)| \leqslant rk \int_K |\varphi|.$$

Now, we can complete the proof of Theorem F1.

One can easily verify, by using Lemma V1, Lemma V2 and the Delta inequality, that condition (b) implies condition (a).

G. Uniquely extremal differentials of Teichmüller type. In this section we will give some applications of the Characterization Theorem using Reich sequences. Using new tools available in infinitesimal cotangent space Q (such as compactness of certain families of holomorphic functions and mean value theorem) we can prove some properties of uniquely extremal dilatation of Teichmüller type.

Let X be a linear space with norm and let M be a subspace of X and let λ be a real bounded linear functional on M. We define the functional $\delta = \delta_\lambda$ on M by $\delta(\varphi) = \|\lambda\| \, \|\varphi\| - \lambda(\varphi)$, $\varphi \in M$. We say that a sequence $\{\varphi_n\}$ from M is weak Hamilton sequence for λ if $\delta_n = \delta(\varphi_n)$ converges to zero.

Suppose now that $X = L^1(\Delta)$ and $M = Q$. Recall that we denote by $Q = Q(\Delta)$ the space of L^1-integrable analytic functions on Δ, and that $\chi \in L^\infty(\Delta)$ is uniquely extremal in infinitesimal class (abbreviated by $\chi \in HBU$) if the linear function $\Lambda_\chi \in Q^*$ induced by χ, $\Lambda_\chi(\varphi) = (\chi, \varphi)$, has a unique norm-preserving extension from Q to a bounded linear functional on $L^1(\Delta)$.

We say that $\chi \in L^\infty(\Delta)$ satisfies Reich condition on a set $S \subset \Delta$ (or, we say that φ_n is Reich sequence for χ on S) if

(1) there is a weak Hamilton sequence φ_n for λ_χ
(2) $\liminf |\varphi_n(z)| > 0$ a.e. in S.

If χ satisfies Reich condition on Δ we will simply say that χ satisfies Reich condition and, also, that φ_n is Reich sequence. Now, we can state an immediate corollary of Characterization Theorem.

Corollary G1. *If χ is uniquely extremal, then χ satisfies Reich condition on its extremal set.*

Corollary G2. *If $|\chi|$ is constant, then χ is uniquely extremal if and only if χ satisfies Reich condition.*

Example G1. If $f(z) = Kx + iy$, $K > 1$, is the affine stretch, defined on a plane domain D, then the Beltrami coefficient μ of f has the form

$$\mu(z) = k\frac{|\varphi_0|}{\varphi_0}, \quad \text{with } \varphi_0 = 1 \text{ and } k = \frac{K-1}{K+1}.$$

In [Re5], Reich has shown that if there exist a sequence φ_n in $Q(D)$ such that: (a) $\varphi_n(z) \to \varphi_0(z)$ for all $z \in D$ and (b) $\delta(\varphi_n) = \|\varphi_n\| - \operatorname{Re}\int_D \varphi_n \to 0$, then $f(z)$ is uniquely extremal.

Also, in [Re5], Reich showed that a sequence $\varphi(z) = e^{-z/n}$ in $Q(D)$ satisfies conditions (a) and (b) for $D = D_\alpha = \{z : y > |x|^\alpha\}$, with $\alpha > 3$, and asked interesting question.

Question G. Whether the conditions (a) and (b) are not only sufficient but also necessary for unique extremality.

Theorem G2 and Corollary G3 (see below) provide an affirmative answer to this question in more general situation.

Concerning the Reich sequences, the next example is interesting.

Example G2. Let $\psi(z) = (1-z)^{-2}$, $\chi = k\frac{|\psi|}{\psi}$, $0 < k < 1$. Then χ is uniquely extremal on Δ.

It is interesting to note that if $t_n \to 1-$ and $\psi_n(z) = (1-t_n z)^{-2}$, then $\delta(\psi_n) \to \pi(1 - \ln 2)$, when $n \to \infty$. Thus, ψ_n is not Reich sequence.

Recall that we say that μ is Teichmüller differential in general sense on Δ if $\mu = k|\psi|/\psi$, where ψ is an analytic function on Δ, which is not identically zero. We say that Reich sequence ψ_n for $\mu = k|\psi|/\psi$ is normalized at a point $z_0 \in G$ if $\psi_n(z_0) \to \psi(z_0) \neq 0$.

The outline of the proof of the next lemma shows how one can use the presence of analytic function in definition of Teichmüller differential to show that each normalized Reich sequence forms a normal family.

Lemma G1. *Let χ be a Teichmüller differential in general sense defined by $\chi = k|\psi|/\psi$, where k is number in $[0,1)$ and ψ is an analytic function on Δ and let ψ_n be normalized Reich sequence for χ at a point z_0 and let $D = D(z_0, r)$ be the disc such that ψ has no zeros on $\overline{D}$. Then ψ_n converges uniformly to ψ on $\overline{D}$.*

Proof. There exist disc D_1 with center at z_0 and of radius $r_1 > r$, such that $\overline{D} \subset D_1$ and that ψ has no zeros on D_1, and positive numbers m and M such that $m \leqslant |\psi(z)| \leqslant M$ for each $z \in D_1$. Let $\varphi_n = \psi_n/\psi$ and

$$\delta_n = \delta_n(D_1) = k\int_{D_1} |\psi_n| - \operatorname{Re}\int_{D_1} k\frac{\psi_n|\psi|}{\psi}. \tag{1}$$

Hence

$$\delta_n \geqslant km\int_{D_1} (|\varphi_n| - \operatorname{Re}\varphi_n).$$

This inequality, with the mean value theorem, shows that φ_n, and therefore ψ_n, form a normal family on D_1. Therefore, there is a subsequence ψ_{n_k} which converges uniformly to ψ_0 on $\overline{D}$. By letting n to infinity in (1) (by D instead of D_1) and using normalization that $\psi_0 = \psi$ at z_0, we conclude that $\psi_0 = \psi$ on D. This actually shows that ψ_n converges uniformly to ψ on D.

The proof of the next result is based on Lemma G1.

Theorem G2. *Let χ be uniquey extremal on Δ and let χ be Teichmüller differential (in general sense) defined by an analytic function φ. Then every normalized Reich sequence φ_n converges uniformly on compact subsets of Δ to φ.*

Corollary G3. [BLMM] *Let χ be Teichmüller dilatation in general sense defined by some analytic function φ in Δ. Then χ is uniquely extremal if and only if there exists Reich sequence φ_n in Q, which uniformly converges on compact subsets of Δ.*

Further developments of the ideas outlined in the proof of Lemma G1 leads to the following results.

Theorem G3. (The first removable singularity Theorem) *Let K be a compact subset of Δ, $G = \Delta \setminus K$ and φ an analytic function on G. Suppose that*

(a) *μ is an extremal dilatation on Δ*

(b) *$\mu = s|\varphi|/\varphi$ on G,*

where s is non-negative measurable function on G. If there exist two positive constants m and M, such that $m \leqslant |\varphi(z)| \leqslant M$, for all $z \in G$, then

(a) *φ has an analytic extension $\tilde{\varphi}$ from G to Δ*

(b) *$\mu = k|\tilde{\varphi}|/\tilde{\varphi}$ a.e. in Δ.*

Theorem G4. (The second removable singularity Theorem) *Let χ be uniquely extremal on Δ and let χ be multiple of nonnegative measurable function and Teichmüller differential defined by analytic function φ on the complement of a compact set $K \subset \Delta$. Then*

(a) *φ has an analytic extension $\tilde{\varphi}$ from G to Δ*

(b) *$\mu = k|\tilde{\varphi}|/\tilde{\varphi}$ a.e. in Δ.*

During author's work with Božin, Lakić and Marković, on the subject concerning unique extremality, we wrote several drafts, in which the proofs of some versions of Theorem G3 and Theorem G4 have been given.

H. Unique extremality and approximation sequences. In this section we briefly discuss some results from [BLMM] and only announce new results (see below Lemma H3, Proposition H3–H4 and Theorem H1–H2).

Let M be a subspace of a normed linear space X and let Λ be a linear functional on M and λ be its real part. Put

$$\overline{\lambda}(x_0) = \inf\{\lambda(y) + \|\lambda\|\|y - x_0\| : y \in M\}$$
$$\underline{\lambda}(x_0) = \sup\{\lambda(x) - \|\lambda\|\|x - x_0\| : x \in M\}.$$

Analysis of the proof of the Hahn-Banach Theorem leads to the following.

Proposition H1. *Linear functional Λ has a unique Hahn-Banach norm-preserving extension from M to X if and only if $\underline{\lambda}(x_0) = \overline{\lambda}(x_0)$, for each $x_0 \in X \setminus M$.*

The details of the proof are left to the reader.

Lemma H1. *Let λ have unique Hahn-Banach norm-preserving extension from M to X. Then for each $x_0 \in X \setminus M$ there exist sequences u_n, $v_n \in M$ such that*
(1) $\lambda(u_n) = \lambda(x_0) + \|\lambda\|\|u_n - x_0\| + o(1)$
(2) $\lambda(v_n) = \lambda(x_0) - \|\lambda\|\|v_n - x_0\| + o(1)$
(3) $\lambda(w_n) = \|\lambda\|(\|x_0 - u_n\| + \|x_0 - v_n\|) + o(1)$, *where $w_n = u_n - v_n$*
(4) *w_n is weak Hamilton sequence for λ.*

We say that λ satisfies unique approximation property at $x_0 \in X \setminus M$ if there exist sequences $\{u_n\}$ and $\{v_n\}$, in M, such that the condition (3) is satisfied.

The following result follows from Lemma 1 and Proposition 1.

Proposition H2. *Let Λ be a bounded linear functional on M and λ be its real part. Then Λ has unique Hahn-Banach norm-preserving extension from M to X if and only if λ satisfies unique approximation property at each $x_0 \in X \setminus M$.*

We say that $\psi_0 \in L^1(\Delta)$ is an extremal vector for linear function $\lambda = \lambda_\chi$, if $\lambda(\psi_0) = k\|\psi_0\|$, where $k = \|\chi\|_\infty$.

For a given $\chi \in L^\infty(\Delta)$, it is convenient to mark the extremal vector $\tilde{\chi}$ defined by $\tilde{\chi} = \overline{\chi}$, on E and $\tilde{\chi} = 0$, on $\Delta \setminus E$.

The further discussion will show that $\tilde{\chi}$ has an important role in characterizations of uniquely extremal dilatation χ.

Lemma H2. *If $\chi \in HBU$, then χ satisfies Reich condition on its extremal set E.*

Proof. Applying the part (1) of Lemma to the function ψ, defined by $\psi = \tilde{\chi}$, we find that there exist a sequence $u_n \in A$ such that

$$\lambda(\psi) + \|\psi - u_n\| + o(1) = \lambda(u_n), \text{ where } \lambda \text{ denotes } \lambda_\chi. \tag{5}$$

Since $\lambda(\psi) = \|\psi\|$ and $\lambda(u_n) \leqslant \|u_n\| \leqslant \|\psi\| + \|u_n - \psi\|$, we obtain, using (5), that

$$\|\psi\| + \|u_n - \psi\| + o(1) = \|u_n\|. \tag{6}$$

Hence, $\lambda(u_n) = \|u_n\| + o(1)$. Thus u_n is weak Hamilton sequence for λ.

Using (6) we conclude that $\|\alpha_n\| \to 0$, $n \to \infty$, where $\alpha_n = |\psi| + |\psi - u_n| - |u_n|$. Hence, there exist subsequence α_{n_k} which converges to zero a.e. on Δ. Since, $|u_n| \geqslant |\psi| - \alpha_n$, it follows that $\lim_{k\to+\infty} \inf |u_{n_k}(z)| \geqslant |\psi(z)|$ a.e. on Δ.

Recall that $E = \{z \in \Delta : |\chi(z)| = k\}$.

Lemma H3. *Let ψ_0 be an extremal vector for λ_χ and $A = \{z \in \Delta : \psi_0(z) \neq 0\}$. Then*

$$\chi(z) = k\frac{|\psi_0(z)|}{\psi_0(z)}, \; z \in A \cap E.$$

Proposition H3. *If $\lambda(\tilde{\chi}) = \underline{\lambda}(\tilde{\chi})$, then χ is uniquely extremal on its extremal set E.*

Let $Q_{\bar{\chi}}$ be the smallest subset of L^1 which contains $Q \cup \{\bar{\chi}\}$. For a given $\mu \in L^\infty$, it is convenient to consider $\lambda = \lambda_\mu$ as a linear functional on L^1, defined by $\lambda_\mu(f) = \mathrm{Re}(\mu, f)$, $f \in L^1$.

Proposition H4. *If μ is a extremal dilatation on Δ and $\psi_0 \in L^1$ an extremal vector, then $\overline{\lambda}(\psi_0) = \lambda(\psi_0)$, where $\lambda = \lambda_\mu$.*

Theorem H1. (The Characterization Theorem II) *Let $\chi \in L^\infty$. The following conditions are equivalent*

(7) *χ is uniquely extremal on its extremal set*

(8) *λ_χ has a unique norm-preserving extension from Q to $Q_{\bar{\chi}}$*

(9) *There exists sequence $\{\varphi_n\}$, with $\varphi_n \in Q$ for all n, such that*

$$\lambda(\varphi_n) = \lambda(\bar{\chi}) + \|\chi\|_* \|\varphi_n - \bar{\chi}\| + o(1)$$

(10) *χ satisfies Reich condition on its extremal set E.*

Theorem H2. *Let $\chi \in L^\infty(\Delta)$ and $k = \|\chi\|_\infty$. If there exist Reich sequence φ_n for χ, on its extremal set E, then*

(11) *$k|\varphi_n|/\varphi_n$ converges to χ a.e. on its extremal set E*

(12) *If two-dimensional Lebesgue measure of E is positive, then*

$$\lim_{n\to\infty} \int_E \chi \frac{|\varphi_n|}{\varphi_n} = k.$$

I. Construction. Before we state next results we introduce a class of sets which is important in our investigation. Let K be a compact subset of $\mathbb{C}$ whose complement is connected in $\mathbb{C}$ and interior is empty. We say that K is a special Mergelyan's set (a special M-set). Motivation for this definition is famous Mergelyan's Theorem.

Mergelyan's Theorem. *If K is a compact set in the plane whose complement is connected, if f is a continuous complex function on K which is holomorphic in the interior of K, and if $\epsilon > 0$, then there exist a polynomial P such that $|f(z) - P(z)| < \epsilon$ for all $z \in K$.*

Note that K need not be connected.

Recall that we denote by $Q = Q(\Delta)$ the space of L^1-integrable analytic functions on Δ, and that $\chi \in L^\infty(\Delta)$ is uniquely extremal in infinitesimal class (abbreviated by $\chi \in HBU$) if the linear function $\Lambda_\chi \in Q^*$ induced by χ, $\Lambda_\chi(\varphi) = (\chi, \varphi)$, has a unique norm-preserving extension from Q to a bounded linear functional on $L^1(\Delta)$.

If $\chi \in HBU$ and $k = \|\chi\|_\infty$, then $\operatorname{ess\,sup} |\chi(z)| = k$ over each open set $G \subset \Delta$, as has been observed by Reich [Re7]. Therefore, the following question (see [Re7] and [S6]) is natural.

Question I. Does $\chi \in HBU$ actually imply that $|\chi(z)| = k$ a.e.?

The next theorem shows that the answer to corresponding question, concerning the more general concept of unique extremal dilation, is negative.

Theorem I1. *Let $K \subset \Delta$ be a compact set of positive measure, whose complement is connected and let $\Omega = \Delta \setminus K$. Then there exist $\chi \in L^\infty(\Delta)$ such that χ is zero on K, χ is uniquely extremal on Ω and $|\chi(z)| = k > 0$ a.e. on Ω.*

Outline of proof. Inductively, we can find a sequence of polynomials $\{P_n\}$ and increasing sequence of compact special M-sets K_n such that

(1) K_n and K are disjunct
(2) complement of $F_n = K \cup K_n$ is connected
(3) $|\bigcup_{n=1}^\infty K_n| = |\Omega|$
(4) $|P_n(z)| < 1/2$ for each $z \in K$ and $|P_n(z)| > 2$ for each $z \in K_n$
(5) $\alpha_n |1 - |P_{n+1}(z)|/P_{n+1}(z)| \leqslant 2^{-n}$, for each $z \in K_{n+1}$, where $\varphi_n = P_1 P_2 \dots P_n$, $\alpha_n = \max\{|\varphi_n(z)| : z \in K_{n+1}\}$
(6) $\int_\Delta |\varphi_n| - \int_{F_{n+1}} |\varphi_n| \leqslant 1/n$, where $F_n = K \cup K_n$.

Let $\chi_n = k|\varphi_n|/\varphi_n$, $0 < k < 1$ and define χ to be zero on K. We leave to the reader to show that $\chi_n(z)$ is a Cauchy sequence a.e. on Ω, that is $\chi(z) = \lim \chi_n(z)$ exists a.e. on Ω; and that φ_n is Reich sequence on Ω for χ.

It is interesting that the following surprisingly simple lemma plays a role in construction unique extremal dilatation with nonconstant modulus.

Lemma Re. *If K is a special Mergelyan's set and ν annihilator of Q in L^∞ such that* $\operatorname{supp} \nu \subset K$, *then $\nu = 0$.*

This lemma was proved by Reich in [Re7] (see also [MM1] and [BLMM]). The following result is an immediate corollary of Theorem I1 and Lemma Re.

Theorem I2 [BLMM]. *Let $K \subset \Delta$ be a special Mergelyan's set of positive measure. Then there exist $\chi \in HBU$ such that $\chi(z) = 0$ in K and $|\chi(z)| = k > 0$ a.e. in $K^c = \Delta \setminus K$.*

We refer to the proof of this result as the construction of uniquely extremal dilatation with nonconstant modulus (shortly the construction).

After writing the final version of this paper, Reich [Re9] has modified the proof of Theorem I2, given in [BLMM], using Runge theorem instead of Mergelyan's theorem.

Further simplifications of the construction has been given by author during his lectures at Scoala Normala Superioara Buchurest (SNSB), 2003–2004 (to appear in [M10]).

Outline of new construction. Recall, if $K \subset \mathbb{C}$ is a compact set and do not separate the plane, we say that K is M-set ("Mergelyan set"); we call K a special M-set if in addition K has empty interior and positive 2-dimensional measure.

Using Runge, we can prove the following result.

Lemma Ma. *Let K be a M-set and G be a Jordan domain such that $\overline{G} \cap K = \emptyset$. For given positive numbers p, q and ε, there exists polynomial Q such that*

(a_1) $|Q| < p$ *on* K; (a_2) $q < |Q|$ *on* $\overline{G}$; (a_3) $|1 - |Q|/Q| < \varepsilon$ *on* $F = K \cup \overline{G}$

Remark: There is an entire function which satisfies the above conditions and in addition has no zeros in $\mathbb{C}$. For suitable Q close to constant functions on K and $\overline{G}$ the entire function $\phi = e^{Q-1}$ satisfies the above conditions.

Let K be a M-set and D be a Jordan domain such that $K \subset D$. Then there exists sequence of Jordan-domains J_n such that

$$(1) \qquad J_n \subset \operatorname{Int} J_{n+1}, \quad \bigcup_1^{\infty} \overline{J_k} = D \setminus K.$$

We say that sequence of Jordan-domains J_n exhaust $D \setminus K$. Inductively, we will find a sequence of polynomials P_n and an increasing sequence of Jordan-domains G_n which exhaust $D \setminus K$ such that:

(b_1) $|P_n| < 1/2$ on K and $|P_n| > 2$ on $\overline{G_n}$

(b_2) $\int_D |\varphi_n| - \int_{F_{n+1}} |\varphi_n| < 1/n$, where $\varphi_n = P_1 \cdots P_n$ and $F_n = K \cup \overline{G_n}$.

(b_3) $\alpha_n |1 - |P_{n+1}|/P_{n+1}| < 2^{-n-1}$, where $\alpha_n = \max\{|\varphi_n(z)| : z \in F_{n+1}\}$.

We call $L_n = D \setminus F_n$ canal. Roughly speaking, by (b_1) and (b_2), we control polynomial φ_n respectively on F_n and L_{n+1}, but we do not control on canal L_n (more precisely on $L_n \setminus L_{n+1}$). Thus, we do not have any estimates of growth of α_n from above. At this point, it seems that it is difficult to overcome this problem. However, we can overcome this problem using (b_3), which has a crucial role. More precisely, applications of Lemma Ma (a_3) shows that there exists polynomial P_{n+1} such that estimate in (b_3) holds. Now, we can modify φ_n on F_{n+1} by means of polynomial P_{n+1}; i.e., we construct the function $\varphi_{n+1} = P_{n+1}\varphi_n$.

The function $|\varphi_n|/\varphi_n$ is defined except on the set Z_n of zeros of polynomial φ_n. Let $Z = \bigcup_1^{\infty} Z_n$. If we define μ_n to be 1 on Z for every $n \geqslant 1$ and $\mu_n = |\varphi_n|/\varphi_n$ on $\mathbb{C} \setminus Z$. Since $\varphi_{n+1} = P_{n+1}\varphi_n$, by b_3) we have $\alpha_n|\mu_{n+1}(z) - \mu_n(z)| < 2^{-n-1}$ for $z \in F_{n+1}$.

Since α_n is obviously increasing sequence, then a standard argument shows that $\mu_n(z)$ is a Cauchy sequence on D; that is $\mu(z) = \lim \mu_n(z)$ exists on D and that $\alpha_n|\mu(z) - \mu_n(z)| < 2^{-n}$ for $z \in F_{n+1}$. Hence, $\|\varphi_n\| = \Lambda_\mu[\varphi_n] + 0(1)$ on D. Since $|\varphi_n(z)| \to \infty$ on $D \setminus (K \cup Z)$, φ_n satisfies Re-condition on $D \setminus K$ and therefore μ is uniquely extremal on $D \setminus K$. Hence, we get

Proposition. *If K is a special M-set and χ measurable function, which is equal μ on $D \setminus K$ and $\|\chi\|_\infty \leqslant 1$, then χ is uniquely extremal on D.*

Note that Theorem I2 can be considered as a corollary of this result.

J. Beltrami equation. Suppose that f has L^1 derivatives in the complex plane $\mathbb{C}$ and that $f(z) \to 0$ as $z \to \infty$. With the notation

$$T\omega = -\frac{1}{\pi} \iint_{\mathbb{C}} \frac{\omega(\zeta)}{\zeta - z} d\xi\, d\eta$$

we then obtain from Green's formula

$$(1) \qquad f = T\overline{\partial} f.$$

For smooth ω with compact support we define the Hilbert transform $H\omega$ of ω by $H\omega = \partial(T\omega)$. By differentiation we obtain an expression for H as principle value

$$(H\omega)(z) = \lim_{\epsilon \to 0_+} -\frac{1}{\pi} \iint_{A_\epsilon} \frac{\omega(\zeta)}{(\zeta - z)^2} d\xi\, d\eta,$$

where $A_\epsilon = \{\zeta : \epsilon < |\zeta| < 1/\epsilon\}$.

Fix $0 < k < 1$, and let $L^\infty(k, R)$ denote the measurable functions on $\mathbb{C}$, bounded by k, and supported in the disc B_R. We let $QC^1(k, R)$ denote the continuous differentiable homeomorphisms f of $\mathbb{C}$ such that $\overline{\partial} f = \mu \partial f$, for some $\mu \in L^\infty(k, R)$, normalized so that $f(z) = z + O(1/z)$, as $z \mapsto \infty$.

Let $f \in QC^1(k, R)$. Then by (1),

$$f(z) - z = T(\overline{\partial} f)(z).$$

Thus, if we set $g = \partial f - 1$ and use $\overline{\partial} f = \mu \partial f$, we obtain

$$g = H(\overline{\partial} f) = H(\mu \partial f) = H(\mu g) + H(\mu).$$

In terms of the operator $H_\mu(g) = H(\mu g)$, $g \in L^p$, we obtain the equation

$$(I - H_\mu) g = H(\mu). \tag{2}$$

If we fix $p = p(k) > 2$ so that $\|H_\mu\| < 1$, then $I - H_\mu$ is invertible. Thus, we can solve the equation (2) for g to obtain

$$g = (I - H_\mu)^{-1} H(\mu) \in L^p. \tag{3}$$

Theorem 1. *Fix $0 < k < 1$, $R > 0$ and $p = p(k) > 2$ as above. For $\mu \in L^\infty(k, R)$, there is a function f on $\mathbb{C}$, normalized so that $f(z) = z + O(\frac{1}{z})$ at ∞, with distribution derivatives satisfying the Beltrami equation $\overline{\partial} f = \mu \partial f$.*

Outline of the proof: Define g by (3) and define $f(z) = z + T(\mu g + \mu)$. Since T is the convolution operator with kernel $1/z$ locally in L^1, f is continuous. Moreover, f is normalized at ∞, and

$$\overline{\partial} f = \mu g + \mu, \quad \partial f = 1 + H(\mu g + \mu) = 1 + g$$

in the sense of distributions, so f satisfies the Beltrami equation.

Acknowledgement. I wish to thank Professor, academician Bogoljub Stanković (the editor of Zbornik radova), for inviting me to write this paper. Special thanks are due to Professor Edgar Reich for useful conversations and his interest in author's papers related to extremal quasiconformal mapping and to M. Knežević for typing this paper. Also, we thank to the technical editor Dragan Blagojevć for showing patience during the preparation of the manuscript.

References

[Ah1] Ahlfors, L. V., *Conformal invariants*, McGraw-Hill Book Company, (1973).

[Ah2] Ahlfors, L. V., *Lectures on Quasiconformal mappings*, Princeton, 1966.

[AMM] Anić, I., Marković, V., Mateljević, M., *Uniformly Bounded Maximal Φ-disks, Bers Spaces and Harmonic Maps*, Proc. Amer. Math. Soc. **128** (2000), 2947–2956.

[AMŠ] Anić, I., Mateljević, M., Šarić, D., *Extremal metrics and modules*, Czech. Math. J. (2000), 52(127), No 2, pp. 225-235.

[BA] Beurling, L., Ahlfors, L., *The boundary correspondence under quasiconformal mappings*, Acta Math. **96** (1956), 125–142.

[BLMM] Božin, V., Lakić, N., Marković, V., Mateljević, M., *The unique extremality*, J. d'Analyse Math. **75** (1998), 299–338.

[BMM] Božin, V., Marković, V., Mateljević, M., *The unique extremality in the tangent space of Teichmüller space*, Proceedings of the International Conference "Generalized Functions-Linear and Nonlinear Problems", Novi Sad, A ug. 31 – Sep. 04, 1996; Integral Transforms and Special Functions **6** (1997), 223–227.

[C] Courant, R., *Dirichlet's principle, conformal mappings and minimal surfaces*, New York: Interscience Publishers, inc. (1950).

[CH] Coron, J. M.,and Helein, F., *Harmonic diffeomorphisms, minimizing harmonic maps and rotational symetry*, Composito Mathematico **69** (1989), 175–228.

[EL] Earle, C., J., Li Zhong, *Extremal quasiconformal mappings in plane domains*, Quasiconformal Mappings and Analysis, A Collection of Papers Honoring F. W. Gehring, Peter Duren, Et Al. Eds., Springer, (1998), 141–157.

[EL1] Eells, J., Lemaire, L., *A report of harmonic maps*, Bull. London Math. Soc. **10** (1978), 109–160.

[EL2] Eells, J., Lemaire, L., *Selected topics in harmonic maps*, 85 pg. AMS Conference Board **50** (1980).

[Ga] Gardiner, F. P., *Teichmüller theory and quadratic differentials*, Wiley-Intersci., New York (1987).

[Ge1] Gehring, F. W., *Rings and quasiconformal mappings in space*, Trans. Amer. Math. Soc. **103** (1962), 383–393.

[Ge2] Gehring, F. W., *Quasiconformal mappings in space*, Bull. Amer. Math. Soc. **69** (1963).

[Ha] Hartman, P., *On homotopic harmonic maps*, Can. J. Math. **19** (1967), 673–687.

[J] Jost, J., *Two dimensional geometric variational problems*, J. Wiley & Sons, 1991, 1–236.

[K] Kühnau, R. *Verzerrungssätze und Koeffizientenbedingungen vom Grunskyschen Typfür quasikonforme Abbildungen.*, Math. Nachr. **48**(1971), 77–105.

[L1] Lehto, O. *Schlicht functions with a quasiconformal extension*, Ann. Acad. Sci. Fenn. AI Math. **500** (1971), 1–10.

[L2] Lehto, O., *Univalent Functions and Teichmüler Spaces*, Springer-Verlag, New York, 1987.

[Lo] Loewner, C., *On the conformal capacity in space*, J. Math. Mech. **8** (1959), 411–414.

[LT1] Li, P., Tam, L. F., *The heat equation and harmonic maps of complete manifolds*, Inv. Math. **105** (1991), 1–46.

[LT2] Li, P., Tam, L. F., *Uniqueness and regularity of proper harmonic maps II*, Indiana U. Math. J. **42** (1993), 593–635.

[LV] Lehto, O., Virtanen, K. I., *Quasiconformal Mapping*, Springer-Verlag, Berlin and New York, 1965.

[M] Mateljević, M., *An extension of the Area Theorem*, Complex Variables **15**, (1990), 155–157.

[M1] Mateljević, M., *An area theorem of Lehto–Kühnau type for harmonic mappings*, (to appear).

[M2] Mateljević, M., *Some geometric question related to harmonic map*, Abstract, Symposium Contemporary Mathematics, Univ. of Belgrade, 18-20 Dec. (1998).

[M3] Mateljević, M., *Dirichlet's principle, uniqueness of harmonic maps and related problems*, Proc. Symp. Contemporary Mathematics, Univ. of Belgrade (1998), 251–267.

[M4] Mateljević, M., *The unique extremality II*, Mathematical Reports 2 (**52**) No. 4, 2000, 503–525.

[M5] Mateljević, M., *Some geometric problems related to harmonic and QC mappings*, International Conference on Geometric Function Theory dedicated to Herbert Grotzsch 1902-1993, Martin Luther University of Halle-Wittenberg, May 27-30, 2000, Abstracts, 12-12.

[M6] Mateljević, M., *A version of Bloch theorem for quasiregular harmonic mappings*, Rev Roum Math Pures Appliq **47** (2002) 5-6, pp 705-707.

[M7] Mateljević, M., *Estimates of the modulus of derivates of harmonic univalent mappings*, Rev Roum Math Pures Appliq **47** (2002) 5-6, pp. 709–711.

[M8] Mateljević, M., Harmonic mappings, trajectory structures and related problems, ABSTRACTS, 5th Congress of Romanian mathematicians, June 22-28, 2003, Pitesti, Romania, pp. 80–81.

[M9] Mateljević, M., *Dirichlet's principle, distortion and related problems for harmonic mappings*, Publ Inst. Math. (Beograd) 75(89) (2004), 147–172

[M10] Mateljević, M., Lectures on quasiconformal mappings, Scoala Normala Superioara Buchurest, (SNSB), 2003-2004.

[MM1] Mateljević, M., Marković, V., *The unique extremal QC mapping and uniqueness of Hahn-Banach extensions*, Mat. Vesnik **48** (1996), 107–112.

[MM2] Marković, V., Mateljević, M., *New version of Grötzsch principle and Reich–Strebel inequality*, Mat. Vesnik **49** (1997), 235–239.

[MM3] Marković, V., Mateljević, M., *New version of the main inequality and uniqueness of harmonic maps*, J. d'Analyse Math. **79** (1999).

[MM4] Mateljević, M., Marković, V., *Reich–Strebel inequality, harmonic mappings and uniquely extremal q.c. mapping*, Abstract, Nevanlinna Colloquium, Lausanne (1997).

[Po] Pommerenke, Ch., *Univalent functions*, Vandenhoeck and Ruprecht, Göttingen (1975).

[Re1] Reich, E., *A generalized Dirichlet integral*, J. d'Analyse Math. **30** (1976), 456–463.

[Re2] Reich, E., *On the variational principle of Gestenhaber and Rauch*, Ann. Acad. Sci. Fenn. **10** (1985), 469–475.

[Re3] Reich, E., *The unique extremality counterexample*, J. d'Analyse Math. **75** (1998), 339–347.

[Re4] Reich, E., *Uniqueness of Hahn-Banach extensions from certain spaces of analytic functions*, Math. Z. **167** (1979), 81–89.

[Re5] Reich, E., *A criterion for unique extremality of Teichmüller mappings*, Indiana Univ. Math. J. **30** (1981), 441–447.

[Re6] Reich, E., *On criteria for unique extremality for Teichmüller mappings*, Ann. Acad. Sci. Fenn **6** (1981), 289–301.

[Re7] Reich, E., *On the uniqueness question for Hahn-Banach extensions from the space of L^1 analytic functions*, Proc. Amer. Math. Soc. **88** (1983), 305–310.

[Re8] Reich, E., *Extremal extensions from the circle to the disc*, Quasiconformal Mappings And Analysis, A Collection Of Papers Honoring F.W.Gehring, Peter Duren Et Al. Eds., Springer, (1998), 321–335.

[Re9] Reich, E., *Extremal Quasiconformal Mapping of the Disk*, Chapter 3 in Geometric function theory, Volume 1, Edited by R. Kühnau, Elsevier, 2002.

[ReS1] Reich, E., Strebel, K., *On quasiconformal mapping which keep the boundary points fixed*, Trans. Amer. Math. Soc. **136** (1969), 211–222.

[ReS2] Reich, E., Strebel, K., *On the Gerstenhaber-Rauch principle*, Israel J. Math. **57**(1) (1987), 89–100.

[Ru] Rudin, W., *Real and complex analysis*, McGraw-Hill 1966.

[Sa] Sampson, J. K., *Some properties and application of harmonic mappings*, Ann. École Norm. Sup. **11** (1978), 211–228.

[Sc] Schoen, R., *The role of harmonic mappings in rigidity and deformation problems*, Lecture Notes in Pure and Appl. Math. **143** (1993), 179–200.

[SY] Schoen, R., Yau, S. T., *Lectures on Harmonic Maps*, Conf. Proc. and Lect. Not. in Geometry and Topology, Vol. II, Inter. Press, (1997).

[S1] Strebel, K., *Zur Frage der Eidentigkeit extremaler quasikonformer Abbildungen des Einheitskreiss II*, Coment. Math. Helv. **39** (1964), 77–89.

[S2] Strebel, K., *Quadratic Differential*, Springer Verlag (1984).

[S3] Strebel, K., *Eine Abschätzung der Länge gewisser Kurven bei quasikonformer Abbildung*, Ann. Acad. Sci. Fenn. **243** (1957), 1–19.

[S4] Strebel, K., *On the existence of extremal Teichmüller mappings*, J. Analyse Math. **30** (1976), 464–480.

[S5] Strebel, K., *On quasiconformal mappings of open Riemann surfaces* Comment. Math. Helv. **53** (1978), 301–321

[S6] Strebel, K., *Extremal quasiconformal mappings*, Resultate der Mathematik **10** (1986), 168–210.

[W] Wan, T., *Constant mean curvature surfaces, harmonic maps and universal Teichmüller space*, J. Diff. Geom. **35** (1992), 643–657.

[We] Wei, H. B., *On the uniqueness problem of harmonic mappings*, Proc. Amer. Math. Soc., **124** (1996), 2337–2341