Mongolian Mathematical Journal

2014, Vol. 18

Ulaanbaatar, 2014
CONTENTS

Sheng Bau., Wenbin Yuan

Edge minimum graphs with given order and bandwidth ........................................... 3

Sheng Bau

Contraction of a closed neighbourhood in bipartite plane graphs .................................... 7

Lutbat Yadamsuren., Won-Joo Hwang

Bayesian Restoration of Wireless Sensor Networks ...................................................... 13

Bayartugs T., Enkhbat R.

On the Maximization of Product of Two Concave Functions ........................................ 23

Bubentchikov A.M., Shcherbakov N.R., Tsyrenova V.B.

Mathematical modelling of forming teeth of a bevel gear with eccentrically-cycloidal gearing ................................................................. 26

Buldaev A.S., Ankhbayar G., Khishegtueva I.-Kh.D.

The fixed point method in polynomial with respect to state problem of parametric optimization ................................................................. 38
Edge minimum graphs with given order and bandwidth

Sheng Bau\textsuperscript{a}, Wenbin Yuan\textsuperscript{b}

\textsuperscript{a}School of Mathematics, University of the Witwatersrand, Johannesburg, South Africa and Institute of Discrete Mathematics, Inner Mongolia University of Nationalities, Tongliao, China
\textsuperscript{b}Center for Discrete Mathematics, Fuzhou University, Fuzhou 350002, China

Abstract

Let $G$ be a finite simple undirected graph with vertex set $V(G)$ and edge set $E(G)$. The order of a graph $G$ is denoted $|G| = |V(G)|$ and the size $|G| = |E(G)|$. A bijection $f : V(G) \rightarrow \{1, 2, \cdots, |G|\}$ is called an ordering of $V(G)$. Define

$$B(f) = \max\{|f(u) - f(v)| : uv \in E(G)\}$$

and

$$B(G) = \min\{B(f) : f \text{ an ordering of } V(G)\}.$$ 

The parameter $B(G)$ is called the bandwidth of $G$.

In this note, we determine all edge minimum graphs with order $p$ and bandwidth $p - 2$.

Keywords: bandwidth, minimum, ordering,

1. Bandwidth of graphs

Let $G$ be a finite simple undirected graph with vertex set $V(G)$ and edge set $E(G)$. The order of a graph $G$ is denoted $|G| = |V(G)|$ and the size $|G| = |E(G)|$. A bijection $f : V(G) \rightarrow \{1, 2, \cdots, |G|\}$ is called an ordering of $V(G)$. Define

$$B(f) = \max\{|f(u) - f(v)| : uv \in E(G)\}$$

and

$$B(G) = \min\{B(f) : f \text{ an ordering of } V(G)\}.$$ 

The parameter $B(G)$ is called the bandwidth of $G$.

The concept of bandwidth was introduced by Harary [3] and originated from bandwidth of a matrix. In 1976, Papadimitriou proved that the bandwidth problem is $NP$-complete [10]. The bandwidth problem is $NP$-complete even for trees with maximum degree 3 [8]. Bandwidth problem has applications in coding theory, interconnection networks, data structures and VLSI design and development [4, 5, 9].

\textit{URL:} sheng.bau@wits.ac.za (Sheng Bau)

The edge minimum bandwidth problem was proposed by software systems design in computer science. This is to find graphs with minimum number of edges which has order $p$ and a given bandwidth. As in [6, 7], the number of edges of such a graph is denoted by $e(p, B)$ where $p = |G|$ and $B = B(G)$. Let
\[ E(p, B) = \{ G : |G| = p, B(G) = B, \|G\| = e(p, B) \}. \]

There are few results on edge minimum bandwidth problem. In this paper, all edge minimum graphs of order $p$ and bandwidth $p - 2$ are determined.

2. Backgrounds

In this section, results relevant to our work will be reviewed.

**Theorem 2.1 (Alavi, McCanna and Erdős [1]; Dutton and Brigham [6]).** For $B \leq \frac{p}{2}$, $e(p, B) = p - 1$, the graph obtained by attaching a path of length $p - 2B$ at a vertex of degree 1 of $K_{1,2B-1}$ is an $e(p, B)$ edge minimum graph. In particular, for $B = \frac{p}{2}$, $K_{1,2B-1}$ is a unique $e(p, B)$ edge minimum graph.

**Theorem 2.2 (Dutton and Brigham [6]).** The complete graph $K_p$ is a unique edge minimum graph of order $p$ and bandwidth $p - 1$. Let $B = p - 2$, $p = 3k + r$ ($k \geq 1, 1 \leq r < 3$), then the complete $(k + 1)$-partite graph $K_{3,\ldots,3,r}$ is an edge minimum graph.

**Theorem 2.3 (Chung [5]).** Let $p = n_1 + \cdots + n_k$, $n_1 = \max\{n_i : 1 \leq i \leq k\}$. Then
\[ B(K_{n_1,n_2,\ldots,n_k}) = p - \left\lceil \frac{n_1 + 1}{2} \right\rceil. \]

For integer $k \geq 1$ let
\[ X = \{x_1, \ldots, x_k\}, \ Y = \{y_1, \ldots, y_k\}. \]
(it is assumed that $X \cap Y = \emptyset$). Define a bipartite graph $H_k$ with
\[ V(H_k) = X \cup Y, \ E(H_k) = \{x_iy_j : i + j > k\}. \]

For illustration, the graphs $H_1, H_2$ and $H_3$ are depicted in Figure 2.

**Theorem 2.4 (Bondy and Murty [2]).** Let $k \leq \frac{p}{2}$. Then $B(G) \geq |G| - k$ if and only if $H_k \not\subseteq G$.

**Theorem 2.5 (Dutton, Brigham and Ingrid [7]).** Let $K_{i,\ldots,i}$ be the complete $m$-partite graph of order $p$ with each part of order $i = \frac{p}{m}$. Then
\[ |E(G)| = \binom{n - i}{2} + (m - 1) \binom{i + 1}{2}. \]
3. Edge minimum graphs

We now determine all edge minimum graphs of order \( p \) and bandwidth \( p - 2 \). It will be agreed that \( K_{1,0} = K_1 \).

**Theorem 3.1.** Let \( p = 3k + r \) (\( k \geq 1, \ 0 \leq r < 3 \)). For each edge minimum \( e(p, p-2) \) graph \( G \), let \( G_i \) be a component of \( G \). Then

1. if \( |G_i| = 3 \) then \( G_i = K_3 \) or \( K_{1,2} \);
2. if \( |G_i| \neq 3 \) then \( G_i = K_{1,r} \);
3. (a) for \( r = 0 \) each component of \( G \) is \( K_3 \); (b) for \( r = 1 \), one component of \( G \) is \( G_{1,i} \) (0 \( \leq i \leq p-4 \), \( i \) is a multiple of 3) and all other components of \( G \) are \( K_3 \); (c) for \( r = 2 \), one component of \( G \) is \( G_{1,i} \) (0 \( \leq i \leq p-4 \), \( i \equiv 1 \) (mod 3)) and and all other components of \( G \) are \( K_3 \).

**Proof:** Let \( G \) be a graph with \( |G| = p \), \( B(G) = p - 2 \) and \( ||G|| \) minimum. Since \( p \geq 4 \), \( \frac{p}{2} \geq 2 \). Then \( B(G) \geq p - 2 \). By Theorem 2.4, \( H_2 \nsubseteq G \).

If \( |G_i| = 3 \), since \( G_i \) is connected, \( G_i = K_3 \) or \( K_{1,2} \). If \( |G_i| < 3 \) then \( G_i = K_i \) for \( i = 1 \) or 2. If \( |G_i| > 3 \), then \( G_i \) is connected and \( H_2 \nsubseteq G_i \). Since \( H_2 = P_4 \) the path of order 4, the diameter of \( G_i \) is less than 3. Now \( G_i \) is acyclic. For if there is a cycle, then \( H_2 \subseteq G_i \). Thus \( G_i \) is either \( K_{1,2} \) or \( K_3 \) for \( |G_i| = 3 \) and is \( K_{1,r} \) for \( |G_i| \geq 4 \). Thus, (1) and (2) follow.

We now prove (3). By Theorem 2.2, for \( r = 1 \) or 2, the complete \((k + 1)\)-partite graph \( K_{3,\ldots,3,r} \) is an edge minimum graph for \( B(G) = p - 2 \). For any edge minimum graph \( G \) with \( B(G) = p - 2 \), \( ||G|| = p - 1 \). Since \( p \geq 4 \), \( H_2 \nsubseteq G \). Hence \( G \) has at least two components. By (1) and (2), each acyclic component of \( G \) has diameter at most 3, that is a star or \( K_1 \). If a component contains a cycle, then it is \( K_3 \). Also since \( ||G|| = p - 1 \), there is at most one component of \( G \) that is \( K_{1,i} \) (1 \( \leq i \leq p - 4 \)). For \( p = 3k + 1 \), let the number of components isomorphic to \( K_3 \) be \( a \). Then \( |K_{1,i}| = 3(k - a) + 1 \). Hence \( i \) divides 3. This proves (a) of (3).

The proof of (b) of (3) is similar.

For \( r = 0 \), by Theorem 2.2, the complete multipartite graph \( K_{3,\ldots,3} \) is an edge minimum graph for \( B(G) = p - 2 \). But

\[
||K_{3,\ldots,3}|| = 3k + 3 = |K_{3,\ldots,3}|.
\]

![Figure 1: Graphs H₁, H₂ and H₃.](image-url)
Hence each component of $\tilde{G}$ is a $K_3$. □

The next corollary presents the edge minimum graphs for $B(G) = p−2$ explicitly. The definitions of join $\vee$ and disjoint union $\cup$ are as given in [2].

**Corollary 3.1.** Let $p = 3k + r$ ($k \geq 1$, $0 \leq r < 3$). Then

1. For $r = 0$, $E(p, p−2) = \{K_{3,...,3}\};$
2. For $r = 1$, $E(p, p−2) = ((K_{3,...,3} \vee K_{p−3j−1}) \cup (K_1 \vee K_{3,...,3}) : 1 \leq j \leq k);$  
3. For $r = 2$, $E(p, p−2) = ((K_{3,...,3} \vee K_{p−3j−1}) \cup (K_1 \vee K_{3,...,3}) : 1 \leq j \leq k).$

**Proof:** The proof is completed by taking complements in Theorem 3.1. □

Contraction of a closed neighbourhood in bipartite plane graphs

Sheng Bau*
School of Mathematics, University of the Witwatersrand,
Johannesburg, South Africa
(sheng.bau@wits.ac.za)

Abstract

In this paper, we consider the contraction of a closed neighbourhood in a connected simple bipartite plane graph and prove a theorem on reductions of such graphs. The reductions produce a relatively large proper minor of the original graph in the same family.

Keywords: bipartite graphs, embedding, facial, plane graphs, sphere

*Work supported by NRF South Africa.
1 Introduction

In this paper, we consider connected simple bipartite plane graphs and prove a theorem on reductions of these graphs. The work of this paper is motivated by [1].

In [1], several theorems on reductions of bipartite graphs and planar bipartite graphs were proved. Results in [1] meet both the requirements as in a theorem of Tutte [4, Theorem 3.2.2, page 58]. These two requirements are: (1) the reduction graph belongs to the same family as the original graph; (2) the reduction graph is a large proper minor of the original graph.

For a graph $G$, we use abbreviations $|G| = |V(G)|$ and $\|G\| = |E(G)|$. For a set $X$, $|X|$ is the cardinality of $X$. Let $G$ be a graph and $A,B \subseteq V(G)$. Denote $[A,B] = \{uv \in E(G) : u \in A, v \in B\}$. If a graph is not specified then $[A,B] = \{uv : u \in A, v \in B, u \neq v\}$. By this convention, if $|A| = m$, $|B| = n$ and $A \cap B = \emptyset$ then $[A,B] \simeq K_{m,n}$. If a subgraph $H$ is written in place of $A$ or of $B$, it is meant $V(H)$. In this paper, for $A \subseteq V(G)$, the subgraph of $G$ induced by $A$ is denoted $G|_A$. The subgraph induced by a subset of edges is understood similarly.

Let $G$ be a connected graph and $A \subseteq V(G)$. If $G \setminus A$ is disconnected, then the union of at least one but not all components of $G \setminus A$ is called a fragment of $G \setminus A$.

Let $G$ be a connected simple plane graph embedded on the sphere. If $p$ is a polygon in $G$ and if $G \setminus V(p)$ is not connected then $p$ is called a separating polygon. Since $G$ is connected and $p$ is separating, for each fragment $F$ of $G \setminus V(p)$, $F \cup p$ is connected. If $L$ is any fragment of $G \setminus V(p)$ and $R = G \setminus (L \cup V(p))$, then we have

1. $\{V(L), V(R), V(p)\}$ is a partition of $V(G)$;
2. $G|_{V(L) \cup V(p)}$ and $G|_{V(R) \cup V(p)}$ are both connected.

An equivalent statement is that there exist connected subgraphs $G_1, G_2 \subseteq G$ such that

1. $G_1 \cup G_2 = G$, and
2. $G_1 \cap G_2 = p$.

The reader is referred to [2] for details of the equivalence stated above.


2 Constructions and minors

A contraction of $G$ is defined to be a partition \{${V_1, V_2, \cdots, V_s}$\} of $V$ such that for each $i = 1, 2, \cdots, s$, the induced subgraph $G|_{V_i}$ is connected. This partition gives rise to a natural mapping from $G$ to a graph $H$, also called a contraction (graph) of $G$. The contraction (graph) $H$ is the graph with

$$V(H) = \{V_1, V_2, \cdots, V_s\}, \quad E(H) = \{V_iV_j : i \neq j, [V_i, V_j] \neq \emptyset\}.$$  

The mapping $f$ is called a contraction (mapping) from $G$ onto $H$, and $G$ is said to be contractible to $H$. Note that, by the definition of the contraction graph $H$, $H$ is a simple graph if the graph is not considered together with an embedding. Note in passing that the contraction of an edge and contraction of a connected subgraph are both special cases of our definition.

A graph $H$ is a minor of $G$, denoted $H \leq G$, if $G$ has a subgraph contractible to $H$. That is, there is a subgraph $K \subseteq G$ and there is a contraction $f : K \to H$. Note that the concept of a minor is a common generalization of the concept of a subgraph and that of a contraction. That $H$ is a minor of $G$ may be presented more clearly by a diagram

$$
\begin{array}{c}
K \\
\eta \\
\downarrow \quad f
\end{array} 
\xymatrix{ 
K \ar[r]^-{\eta} \ar[d]_f & G \\
H \ar[ur]^\mu & \}
$$

where $\eta : K \to G$ is a subgraph inclusion (or embedding), $f : K \to H$ is a contraction. The mapping (relation) $\mu$ a minor inclusion $H \leq G$.

The graphs in this paper are embedded on the sphere. We denote the topological sphere by $S$. A Jordan curve on $S$ is the image of the unit circle under a continuous mapping into $S$. For a graph $G$ embedded on $S$, a Jordan polygon is just a polygon in $G$. The topological image of a Jordan polygon in $G$ is a Jordan curve on $S$. If $C$ is a Jordan curve on $S$, then the topological space $S \setminus C$ has exactly two components each is homeomorphic to an open disk. Since $G$ is embedded on $S$, each face boundary $C$ is a Jordan polygon such that one of the two components of $S \setminus C$ does not intersect $G$ (i.e., disjoint from $V(G) \cup E(G)$), which may be called the interior of the face. A slight advantage of working with the sphere instead of the plane is where loose talks such as “the outside face” or “the infinite face” is avoided.

For an explanation of my preference of the terminology such as “trilateral” and “quadrilateral” the reader is referred to [2].
3 Contraction of a closed neighbourhood

In this section we make a correction to the proof of a theorem in [1], where the proof was not complete. As in graph theory literature, \(\delta(G)\) denotes the minimum degree of a graph \(G\). That is, \(\delta(G) = \min\{d(u) : u \in V(G)\}\).

**Lemma 3.1.** Let \(G\) be a connected bipartite plane graph. For \(k \geq 1\), suppose that \(G\) has \(n_k\) vertices of degree \(k\). Then

\[
3n_1 + 2n_2 + n_3 \geq 8.
\]

In particular, \(\delta(G) \leq 3\).

**Proof.** Let \(G\) be a connected bipartite plane graph. Denote by \(n_k\) the number of vertices of degree \(k\), by \(p_i\) the number of faces of size \(i\), by \(\nu(G)\), \(\epsilon(G)\) and \(\phi(G)\), respectively, the number of vertices, the number of edges and the number of faces of \(G\) in the embedding. Then the polyhedral formula of Euler is

\[
\nu(G) - \epsilon(G) + \phi(G) = 2. \tag{1}
\]

Since \(G\) is a bipartite graph each \(i\) is even. For each face of size \(i \geq 6\) quadrangulate the face by introducing \(\frac{i}{2} - 2\) edges joining pairs of vertices on the boundary of the face. Hence

\[
4\phi = 2\epsilon - \sum_{i \equiv 0 \pmod{2}} (i - 4)p_i. \tag{2}
\]

We have also

\[
\nu = \sum_{k \geq 1} n_k, \quad 2\epsilon = \sum_{k \geq 1} kn_k.
\]

Hence by (1) and (2) we have

\[
4 \sum_{k \geq 1} n_k - 2 \sum_{k \geq 1} kn_k + \sum_{k \geq 1} kn_k - \sum_{i \equiv 0 \pmod{2}} (i - 4)p_i = 8.
\]

Since \(G\) is a bipartite graph, \(G\) has no face of odd size. In particular, \(G\) has no trilateral face. Hence

\[
\sum_{i \equiv 0 \pmod{2}} (i - 4)p_i \geq 0,
\]

where the sum is over even integers \(i \geq 4\). Hence

\[
3n_1 + 2n_2 + n_3 = 8 + \sum_{k \geq 4} (k - 4)n_k + \sum_{i \equiv 0 \pmod{2}} (i - 4)p_i \geq 8. \tag{3}
\]

The proof of the lemma is complete. \(\square\)
We now provide a proof of a theorem in [1] where the proof was not complete. Note that as in [1], $\bar{N}(x) = \{x\} \cup N(x)$ denotes the closed neighbourhood of $x \in V(G)$. Note that the statement of the theorem is also slightly different from that in [1] while the result is essentially the same.

**Theorem 3.2 ([1, Theorem 2.2]).** Let $G \neq K_2$ be a connected bipartite plane graph. Then there exists $x \in V(G)$ with $d(x) = \delta(G) \leq 3$ such that

$$H = G/G_{\bar{N}(x)}$$

is a connected bipartite plane graph.

**Proof.** Let $G$ be a connected bipartite plane graph. As in [1], the contraction of the closed neighbourhood $\bar{N}(x)$ of a vertex $x \in V(G)$ of minimum degree (by Lemma 3.1, $\delta(G) \leq 3$) yields a graph $H = G/G_{\bar{N}(x)}$ that is connected and bipartite.

Since the contraction of $\bar{N}(x)$ in $G$ is within the surface of the sphere, $H$ is a plane graph embedded in the sphere.

It is needed to prove that $H$ has no loop or multiple edges. This is where the proof was not complete in [1].

We now prove that $H$ has no loop and no multiple edge. Since $d(x) = \delta(G) \leq 3$, we have $d(x) \in \{1, 2, 3\}$. If $d(x) = 1$ then no loop or multiple edges will result in $H$.

Hence $d(x) = 2$ or $3$ and the contraction of $\bar{N}(x)$ contracts two edges which produces multiple edges in the contraction graph. Hence some two edges incident with $x$ are edges of a separating quadrilateral in $G$.

Since $G$ is a bipartite graph it has no polygon of odd size, and in particular, face of odd size. Hence $G$ has no trilateral and hence $G/G_{\bar{N}(x)}$ has no loop.

Assume that for every bipartite plane graph $H$ (embedded on the sphere) with $|H| \geq 4$ and $|H| < |G|$, there exists $x \in V(H)$ with $d(x) = \delta(H) \leq 3$ such that $H/H_{\bar{N}(x)}$ is a bipartite plane simple graph. We refer to this assumption as the inductive hypothesis.

If multiple edges occur in the contraction graph $G/G_{\bar{N}(x)}$ as a polygon of size 2 which is facial, then the multiple edges are replaced by a single edge by our definition of a contraction. Therefore, if $G/G_{\bar{N}(x)}$ has multiple edges then there exists a Jordan polygon $C$ with $\|C\| = 2$ such that in the embedding both components of $S \setminus C$ intersect $G_{\bar{N}(x)}$ and hence intersect $G$. The polygon $C$ corresponds to a polygon $p$ in $G$ with $\|p\| = 4$, whose image under the contraction mapping is $C$. Hence $p$ is a separating quadrilateral of $G$. As argued above, we may also assume that $p$ is not facial.

Let $L$ be a fragment of $G \setminus V(p)$ of minimum order. (This statement will be referred to as the choice of $p$ and $L$ in the sequel.) Let $R = G \setminus (L \cup p)$. Denote
\[ G_1 = G|_{L_p} \text{ and } G_2 = G|_{R_p}. \] Then by the choice of \( p \) and \( L \), \(|G_1| \leq |G_2| \) and \( G_1 \) has no separating quadrilateral. Since \( G_1, G_2 \subset G \), \( \delta(G_1) \leq \delta(G) \). By the inductive hypothesis, there exists \( z \in V(G_1) \) with \( d_{G_1}(z) = \delta(G_1) \leq 3 \) such that \( G_1/G_1|_{N(z)} \) is a bipartite plane simple graph.

If \( z \notin V(p) \) then since \( G_1 \) has no separating quadrilateral, \( H = G/G|_{N(z)} \) is simple and the conclusion of the theorem is true.

Hence suppose that \( z \in V(p) \). Since \( d_{G}(z) = 2 \) or \( 3 \), let \( L' = L \setminus \{z\} \) and \( R' = R \cup \{z\} \) then \( p \) is still a separating quadrilateral in \( G \) and \(|L'| < |L|\). This contradicts the choice of \( p \) and \( L \).

This completes the proof. \( \square \)

Every connected bipartite plane graph of order at least 3 has a large proper minor that is also a connected bipartite plane graph.

**Corollary 3.3.** Let \( G \neq K_2 \) be a connected bipartite plane graph. Then there exists a connected bipartite plane graph \( H \) such that \( H \) is a minor of \( G \) and

\[ |G| - 3 \leq |H| \leq |G| - 1. \]

**References**


Bayesian Restoration of Wireless Sensor Networks

Lutbat Yadamsuren, Won-Joo Hwang
Dept of Information and Communication System, Inje University
lutubatu@yahoo.com

Abstract
We examine an analogy between statistical mechanics system and wireless sensor network (WSN). The energy function of physical system determines its Gibbs distribution (GD). The analogy gives a probabilistic graphical model for the WSN which characterizes jointly the spatial dependence of a node and a node-to-node connection with minimal energy. We have shown that our proposed model is a GD. This model enables the stochastic relaxation that allows nodes to configure themselves such that more survivable based on information from neighbors upon network degradation. Simulation results demonstrate significant performance improvements.

1 Introduction
WSNs can be characterized by node positions and node-to-node connections. The survivable network configuration is for network nodes to independently adapt node positions and node-to-node connections to support end-to-end connections between nodes, and finding more survivable network model is still a hot topic. Locality, randomness, spatial dependency and energy consumption are main characteristics of network configuration management of infrastructureless wireless networks such as WSN. For example, nodes can either fail locally or move away from a geographical proximity. Local adaptation is thus desirable for preventing continuous network-wide connection. Furthermore, local adaptation is often unavoidable for WSNs where nodes have local information only from their neighbors. Local information available at a node includes locations of neighboring nodes and communication activities of their adjacent links. Such local information has random properties due to uncertainty of user behavior and incomplete information. The corresponding network variables are spatially dependent due to either constraints on physical and logical configurations or information exchange among neighbors.

In this paper, our model should take locality, randomness and spatial dependency into consideration and allow nodes to self-adapt local configurations, resulting in a network-wide optimal configuration with minimal energy consumption, collectively, in turn it increases network lifetime. Here we adopt a Bayesian approach, and introduce a stochastic model for the actual network condition, based on the GD, and we use an existing restoration stochastic relax-
ation algorithm [3] for computing the maximum a posteriori (MAP) restoration of the actual network via local information. This algorithm is highly parallel and exploits the equivalence between GDs and Markov random field (MRF).

2 Related works

Numerous distributed algorithms and protocols have been developed for topology formation [11], self-organizing sensor networks [2], and p2p self-stabilizing networks (see [6]). These distributed algorithms are deterministic and do not involve random variables. Baras et al. [1] develops a probabilistic model with spatial Markov assumptions to characterize randomness in node positions. Meshkova et al. [8] applied simulated annealing to learn parameters of graphical models for cross-layer optimization. These randomized algorithms assume Markovian spatial dependence that warrants the optimality of the distributed algorithms. Madan et al. [7] show that the general SINR based interference models are non-Markovian. There exists spatial dependence with far interferers when aggregated interference from far interferers is not negligible [7]. Kauffmann et al. [5] develops a GD for measurement based self-organization for 802.11 wireless access networks, where each node is aware of all the other nodes’ decisions via base stations. Different from access networks, individual nodes in wireless networks often do not have complete information of the entire network. Sung-eok Jeon et al. [4] studied the near-optimal versus the complexity of distributed configuration management for wireless networks and they have used two-layered MRF. All these works guide us to the Bayesian model which takes locality, randomness and spatial dependency into consideration and allow nodes to self-adapt local configurations, resulting in a network wide optimal configuration with minimal energy consumption.

3 System model

In this paper, we consider the network upon a situation with a failure as a pair $X = (F, L)$, where $F$ is the node positions and $L$ denotes link activities.

3.1 MRF and GD

Let $S = \{s_1, s_2, \ldots, s_N\}$ be a set of sites and let $G = \{G_s, s \in S\}$ be a neighborhood system for $S$, meaning any collection of subsets of $S$ for which 1) $s \in G_s$ and 2) $s \in G_r \Leftrightarrow r \in G_s$. Obviously, $G_s$ is the set of neighbors of $s$ and the pair $\{S, G\}$ is a graph in the usual way. A subset $C \subseteq S$ is a clique if every pair of distinct sites in $C$ are neighbors; $C$ denotes the set of cliques. The index set $S$ has enough geometric structure to even define any kind of WSNs. Let $X = \{X_s, s \in S\}$ denote any family of random variables indexed by $S$. For simplicity, we can assume a common state space, say $\Lambda$, so that $X_s \in \Lambda$ for all $s$; the extension to site dependent state spaces, appropriate when $S$ consists
of both node and link sites, is entirely straightforward. Let \( \Omega \) be the set of all possible configurations:

\[
\Omega = \{ \omega = (x_{s_1}, x_{s_2}, \ldots, x_{s_N}) | x_{s_i} \in \Lambda, 1 \leq i \leq N \}.
\]

\( X \) is an MRF with respect to \( G \) if

\[
P(X = \omega) > 0 \text{ for all } \omega \in G; \\
P(X_s = x_s | X_r = x_r, r = s) = P(X_s = x_s | X_r = x_r, r \in G_s)
\]

for every \( s \in S \) and \( (x_{s_1}, x_{s_2}, \ldots, x_{s_N}) \in \Omega \).

The left-hand side of (2) is called the local characteristics of the MRF and it turns out that the joint probability distribution \( P(X = \omega) \) of any process satisfying (1) is uniquely determined by these conditional probabilities.

A GD relative to \( \{S, G\} \) is a probability measure \( \pi \) on \( \Omega \) with the following representation:

\[
\pi(\omega) = e^{-U(\omega)/T} / Z,
\]

where \( Z \) and \( T \) are constants and \( U \), called the energy function, is of the form

\[
U(\omega) = \sum_{C \in C} V_C(\omega).
\]

Each \( V_C \) is a function on \( \Omega \) with the property that \( V_C(\omega) \) depends only on those coordinates \( x_s \) of \( \omega \) for which \( s \in C \). Such a family \( V_C, C \in C \) is called a potential. \( Z \) is the normalizing constant: \( Z = \sum_\omega e^{-U(\omega)/T} \) and is called the partition function. \( T \) controls the degree of peaking in the density \( \pi \) and choosing \( T \) small clarifies the mode(s), making them easier to find by sampling; this is the principle of annealing, and will be applied to the posterior distribution (PD) \( \pi(f,l) = P(F = f, L = l | G = g) \) in order to find the MAP estimate. We can show that \( \pi(f,l) \) is GD and identify the energy in terms of those for the priors.

**Theorem 1.** Let \( G \) be a neighborhood system. Then \( X \) is an MRF with respect to \( G \) iff \( \pi(\omega) = P(X = \omega) \) is a GD with respect to \( G \).

This equivalence provides us with a simple, practical way of specifying MRF’s. The proof can be found in many of references.

### 3.2 Posterior distribution

We now turn to the PD \( \pi(f,l) \) of the network given the current information \( G \). The configuration space is the set of all pairs \( \omega = (f,l) \), where \( f \) is the allowable node position and \( l \) is the coded link activities.

Assume \( X \) is an MRF relative to \( \{S, G\} \) with corresponding energy function \( U \) and potentials \( \{V_C\} : \)

\[
P(F = f, L = l) = e^{-U(f,l)/T} / Z,
\]
\[ U(f, l) = \sum_C V_C(f, l). \] (6)

For convenience, take \( T = 1 \). We denote the random shift as \( \mathcal{N} \), due to channel conditions, mobility and environmental changes and each node has some random shift, i.e., \( G = F + \mathcal{N} \), where \( \mathcal{N} \) has Gaussian distribution with a mean \( \mu \) and variance \( \sigma^2 \) and it is independent of \( X \).

Let the collection \( \{ \mathcal{F}_s \}, s \in S_1 \) be a neighborhood system in \( S_1 \), where \( S_1 \) is a set of node sites and \( \mathcal{F}^2 \) denotes the second order system, i.e., \( \mathcal{F}^2_s = \bigcup_{r \in \mathcal{F}_s} \mathcal{F}_r, \quad s \in S_1 \). \( \{ \mathcal{F}^2_s \setminus \{ s \}, s \in S_1 \} \) is also a neighborhood system. The set \( \mathcal{G}^P = \{ \mathcal{G}^P_s, s \in S \} \), where

\[
\mathcal{G}^P_s = \begin{cases} 
\mathcal{G}_s, & \text{if } s \in S_2; \\
\mathcal{G}_s \cup \mathcal{F}^2_s \setminus \{ s \}, & \text{if } s \in S_1,
\end{cases}
\]

where \( S_2 \) is a set of link sites. The \( P \) stands for posterior; some thoughts show that \( \mathcal{G}^P_s \) is neighborhood system on \( S \).

Let the collection \( \{ \mathcal{F}_s, s \in S_1 \} \) be a neighborhood system in \( S_1 \), where \( S_1 \) is a set of node sites and \( \mathcal{F}^2 \) denotes the second order system, i.e., \( \mathcal{F}^2_s = \bigcup_{r \in \mathcal{F}_s} \mathcal{F}_r, \quad s \in S_1 \). \( \{ \mathcal{F}^2_s \setminus \{ s \}, s \in S_1 \} \) is also a neighborhood system. The set \( \mathcal{G}^P = \{ \mathcal{G}^P_s, s \in S \} \), where

\[
\mathcal{G}^P_s = \begin{cases} 
\mathcal{G}_s, & \text{if } s \in S_2; \\
\mathcal{G}_s \cup \mathcal{F}^2_s \setminus \{ s \}, & \text{if } s \in S_1,
\end{cases}
\]

where \( S_2 \) is a set of link sites. The \( P \) stands for posterior; some thoughts show that \( \mathcal{G}^P_s \) is neighborhood system on \( S \).

Let \( || \cdot || \) denote the usual norm in \( R^N \).

Now we are ready to formulate the following assertion.

**Theorem 2.** For each \( g \) fixed, \( P(X = \omega | G = g) \) is a GD over \( \{ S, \mathcal{G}^P \} \) with energy function

\[
U^P(f, l) = U(f, l) + \frac{||\mu - \mathcal{N}||^2}{2\sigma^2}. \] (7)

**Proof:** Using standard results about regular conditional expectations, we can assume that

\[
P(X = \omega | G = g) = \frac{P(X = \omega) P(G = g | X = \omega)}{P(G = g)} \] (8)

for all \( \omega = (f, l) \), for each \( g \). Since \( P(G = g) \) is a constant and \( P(X = \omega) = e^{-U(\omega)/Z} \), the key term is

\[
P(G = g | X = \omega) = P(g = f + \mathcal{N} | F = f, L = l) = P(\mathcal{N} = g + f | F = f, L = l) = P(\mathcal{N})
\]

(since \( \mathcal{N} \) is independent of \( F \) and \( L \))

\[
= (2\pi\sigma^2)^{-N^2/2} \exp\left(\frac{1}{2\sigma^2} ||m - \mathcal{N}||^2\right)
\]

Collecting constants, we have, from (8),

\[
P(X = \omega | G = g) = e^{-U^P(\omega)/Z^P}
\]

for \( U^P(\omega) \) as in (7); \( Z^P \) is the usual normalizing constant. Take any \( s \in S_2 \), the local characteristics at \( s \) for the PD are, by (7),

\[
P(L_s = l_s | L_r = l_r, r \neq s, r \in S_2, F = f, G = g) = e^{-U^P(f, l)} \sum_{l_s} e^{-U^P(f, l)} = e^{-U(f, l)} \sum_{l_s} e^{-U(f, l)}
\]
where the sum extends over all possible values of $L_s$. Hence $G_s^P = G_s$.

For $s \in S_1$, the term in (7) involving $N$ does not cancel out. Now $N = \{\eta_s|s \in S_1\}$ and let us denote the dependencies by writing $\eta_s = \eta_s(g_s-f_t,t \in F_s)$. Then

$$P(F_s = f_s|F_r = f_r, r \neq s, r \in S_1, L = l, G = g) = e^{-U^P(f,l)}/\sum_{L_s} e^{-U^P(f,l)};$$

$$U^P(f,l) = U(f,l) + \sum_{r \in S_1} (\eta_r - \mu)^2/2\sigma^2. \quad (9)$$

Decompose $U^P$ as follows:

$$U^P(f,l) = \sum_{C:s \in C} V_C(f,l) + \frac{1}{2\sigma^2} \sum_{r:s \in F_r} (\eta_r(g_r - f_t, t \in F_r) - \mu)^2 + \sum_{C:s \notin C} V_C(f,l) + \frac{1}{2\sigma^2} \sum_{r:s \notin F_r} (\eta_r(g_r - f_t, t \in F_r) - \mu)^2.$$

Since the last two terms do not involve $f_s$, the ratio in (9) depends only on the first two terms above. The first term depends only on coordinates of $(f,l)$ for sites in $G_s (s \in C \Rightarrow C \subseteq G_s)$ and the second term only on sites in $\bigcup_{r:s \in F_r} F_r = \bigcup_{r:F_r} F_r = \bigcup_{r:s \in F_r} F_r^2$.

Hence, $G_s^P = G_s \cup F_s^2 \setminus \{s\}$, as asserted in the theorem □

4 Problem formulation

4.1 Assumptions and notation

All nodes share a common frequency channel and the wireless channel follows a path-loss model with power attenuation factor $2 \leq \alpha \leq 5$. For simplicity, multipath fading and shadowing are not considered in this work. Node $i$, $1 \leq i \leq N$ transmits with power $P_i$, where $0 \leq P_i \leq P_{\text{max}}$. $P_{\text{max}}$ is the maximum transmission power, and $N$ is the number of nodes in the WSN.

Let $\text{SINR}_{\text{th}}$ be a given threshold for the SINR requirement. Node $i$ can successfully transmit to $j$ node when the SINR requirement is satisfied, i.e.,

$$\text{SINR}_{(i,j)} = \frac{P_i d_{(i,j)}^{\alpha}}{N_p + \sum_{(m,n) \neq (i,j)} P_m d_{(m,j)}^{\alpha}} \geq \text{SINR}_{\text{th}},$$

where $\text{SINR}_{(i,j)}$ is the SINR of active link $(i,j)$, $d_{(i,j)}$ is the distance between nodes $i$, $j$ and $N_p$ is noise power. Let $f_i$ and $\bar{f}_i$ be an actual and an expected location of node, for $1 \leq i \leq N$. Let $l_{(i,j)}$ denote the link-activity of link $(i,j)$, where $l_{(i,j)} = 1$ if node $i$ is transmitting to node $j$; and $l_{(i,j)} = 0$, otherwise.
4.2 Formulation

The GD for local configuration with optimal node positions and link-activities that minimizes the composite energy consumption function, in turn it increases network lifetime. Total energy is found as:

\[ U(f, l) = U(f | l) + U(l) \]

where

\[ U(f | l) = \sum_{(i,j)} \alpha_{i,j} l(i,j) - \sum_{(i,j)} P_j l(i,j) + \beta (\text{SINR}_{i,j} - \text{SINR}_{th})^2 l(i,j) \]

\[ U(l) = \sum_{(i,j)} \sum_{(m,n) \in N(i,j)} \alpha_{i,j,(m,n)} l(i,j) l(m,n) \]

The coefficients of total energy are found in [4]:

\[ \alpha_{i,j} = -P_i d_{i,j}^{-\alpha} + \beta (-P_i d_{i,j}^{-\alpha} - \text{SINR}_{th} N_P)^2 \]

\[ \alpha_{i,j,(m,n)} = 2 \sqrt{P_i P_m d_{i,j}^{-\alpha} d_{m,j}^{-\alpha}} - P_m d_{m,j}^{-\alpha} + \beta \text{SINR}_{th}^2 P_m^2 d_{m,j}^{-2\alpha} - 2\beta (-P_i d_{i,j}^{-\alpha} - \text{SINR}_{th} N_P) \text{SINR}_{th} P_m d_{m,j}^{-\alpha} \]

where \( \beta > 0 \) is a weighting factor. Intuitively, \( \alpha_{i,j} \) corresponds to the increased power of node \( j \) when link \( (i, j) \) becomes active, \( \alpha_{i,j,(m,n)} \) relates to the interference experienced by \( (i, j) \) resulting from neighboring active link \( (m, n) \).

By the Theorem 2 total energy function takes the following correction because of randomness

\[ \frac{1}{2\sigma^2} \| \mathcal{N} - \mu \|^2 = \sum_i \frac{(n_i - \mu_i)^2}{2\sigma^2} \]

where \( \sigma \) is assumed to be the same for all nodes for simplicity. If we take \( \mu = 0 \) then

\[ \| \mathcal{N} - \mu \|^2 = \sum_i (f_i - \tilde{f}_i)^2 \]

4.3 Distributed algorithm

We adopt a Bayesian kind of approach based on stochastic relaxation and annealing, for computing the MAP estimate of the WSN. This algorithm is highly parallel and exploits the equivalence between GDs and MRF and converges to a minima asymptotically with probability 1.

Let us imagine a simple processor placed at each site \( s \) of the graph and the connectivity relation among the processors: the processor at \( s \) is connected to each processor for the sites in \( G_s \). The state of the machine evolves by discrete changes. At time \( t \), the state of the processor at site \( s \) is a random variable \( X_s(t) \) with values in \( \Lambda_s \). The total configuration is \( X(t) = (X_{s_1}(t), \ldots, X_{s_N}(t)) \), which evolves due to state changes of the individual processors. The starting configuration, \( X(0) \), is arbitrary. At each epoch, only one site undergoes a
(possible) change, so that $X(t - 1)$ and $X(t)$ can differ in at most one coordinate. Let $n_1, n_2, \ldots$ be the sequence in which the sites are visited for replacement; thus, $n_t \in S$ and $X_{s_i}(t) = X_{s_i}(t - 1)$, $i \neq n_t$. Each processor is programmed to follow the same algorithm: at time $t$, a sample is drawn from the local characteristics of $\pi$ for $s = n_t$ and $\omega = X(t - 1)$.

The evolution $X(t - 1) \rightarrow X(t)$ in mathematical terms,

$$P(X_s(t) = x_s, s \in S) = \pi(X_{n_t} = x_{n_t} | X_{n_s} = x_{n_s}, s = n_t) P(X_s(t - 1) = x_s, s = n_t),$$

where $\pi$ is the GD. If we indicate the dependence of $\pi$ on $T$ by writing $\pi_T$, and let $T(t)$ denote temperature at stage $t$. The procedure generates a different process $X(t)$, $t = 1, 2, \ldots$ such that

$$P(X_s(t) = x_s, s \in S) = \pi_T(t)(X_{n_t} = x_{n_t} | X_{n_s} = x_{n_s}, s = n_t) P(X_s(t - 1) = x_s, s = n_t),$$

where annealing schedule $T(t) = T_0 / \log(1 + t)$, $1 \leq k \leq \infty$ at $t$th iteration. Note that the amount of time required for one complete update of entire system is independent of the number of sites.

5 Simulation result and evaluation

In a multihop wireless network, an important requirement of such network is strong connectivity but our distributed algorithm does not guarantee any strong connectivity. Besides strong connectivity, an important design metric is energy efficiency. As it directly impact the network lifetime. The work, Takagi et al. [10] tries to optimize energy consumption and maximize network throughput the same time and their algorithm does not guarantee strong connectivity.

Our proposed model minimizes the interference and the distributed algorithm also enhancing throughput. Ramanathan et al. [9] has considered optimizing for the minmax transmission power in centralized algorithms, however their distributed heuristic algorithms do not guarantee strong connectivity. We do our comparison only with [10]. We refer to their algorithm as $T&K$, our basic model algorithm (the distributed algorithm) as $MaxMar$. As a reference, we also compare with the no topology control case where each node always uses the maximum transmission radius for broadcasting a packet ($MaxPower$).

5.1 Simulation environment

The distributed algorithm is implemented in ns-2, using the wireless extension that already developed. Our simulation is done for a network of 100 nodes with WaveLAN-I radios. The nodes are deployed uniformly at random in a rectangular region of $1.5 \times 1.5$ km$^2$. There has been some work on realistic topology generation. Therefore, WSNs are placed automatically. We assume that for the wireless channel, it has attenuation $\alpha = 4$ and the distributed algorithm based on local decisions can satisfy the SINR constraint with $\text{SINR}_{\text{th}} = 20$. 

19
5.2 Network Performance Analysis

We would like to measure the network performance using two different algorithms in [10] and MaxPower. We care about network lifetime in the WSNs environment. We measure the network lifetime as the number of nodes still alive over time.

As can be seen from Fig 1, the distributed MaxMar algorithm performs as good as the T&K algorithm, while using only local information. They both perform significantly better than MaxPower.

Figure 2 presents the evolution the mean potential delay and the improvement observed through the introduction of the proposed algorithm. We also collected throughput statistics at the end of our simulation. The MaxMar algorithm and the T&K algorithm achieve 1.7 times the throughput of the MaxPower. The throughput statistics show that it is undesirable to transmit over large radius will increase energy consumption and also cause unnecessary interference. Increased interference will result in decreased throughput.

6 Conclusion

We made an analogy between statistical mechanics system and WSN. The presence and direction of active links and sensor node positions are assumed as states molecules or atoms in a physical systems. An energy function of the physical system determines its GD. The GD and MRF equivalence, due to this equivalence the energy function determines an MRF network model. For any cause of network degradation, including node and link failure, insufficient power, and node random shift, the PD is an MRF with a structure the same with the degraded WSN. By the analogy, the PD gives another physical system. We have found PD of the degraded WSN, in order to restore the WSN. We experiment with simple WSN, for which restorations are obtained.
Figure 2: Average potential delay per node seen over time for a static topology: with different algorithm

References
On the Maximization of Product of Two Concave Functions

T.B. Ayartugs∗ and R.E. Enkhbat†
∗University of Science and Technology, Mongolia
†National University of Mongolia

Abstract. We consider a problem of maximizing the product of two concave functions. The problem belongs to a class of nonconvex optimization. We propose new global optimality conditions for the problem by applying a result in [1].

1 Introduction

Optimization of a product of convex or concave functions plays important roles not only in optimization theory but also in engineering and economics. The problem is not convex. Therefore, a global solution over a convex feasible set may not be found by standard optimization techniques. In general, multiplicative programming problem can be solved by using an outer approximation algorithm [6].

In the most literature [3–7], so called convex multiplicative programming problem which is a minimization of a product of several convex functions has been studied.

min f(x) = \prod_{j=1}^{p} f_j,
subject to g_j(x) ⩽ 0, j = 1, 2, ..., m,
where f_j: \mathbb{R}^n \rightarrow \mathbb{R}, j = 1, 2, ..., p and g_i: \mathbb{R}^n \rightarrow \mathbb{R}, j = 1, 2, ..., m are convex functions.

The following result is well known [5] for solving problem (1.1)-(1.2). First, the following master problem is solved.

min F(x, λ) = \prod_{j=1}^{p} λ_j f_j,
subject to g_j(x) ⩽ 0, j = 1, 2, ..., m,
p\prod_{j=1}^{p} λ_j ⩾ 1, λ = (λ_1, λ_2, ..., λ_m) ⩾ 0.

Lemma 1.1 [5] Let (x∗, λ∗) be an optimal solution of (1.3)-(1.4). Then x∗ is optimal to (1.1)-(1.2).

Minimization of product of two convex functions was studied in [2]. Unlike this, in this paper we consider a problem of maximizing a product of two concave functions.
On the Maximization of Product of Two Concave Functions

T.Bayartugs* and R.Enkhbat†

* University of Science and Technology, Mongolia
† National University of Mongolia

Abstract. We consider a problem of maximizing the product of two concave functions. The problem belongs to a class of nonconvex optimization. We propose new global optimality conditions for the problem by applying a result in [1].

1 Introduction

Optimization of a product of convex or concave functions plays important roles not only in optimization theory but also in engineering and economics. The problem is not convex. Therefore, a global solution over a convex feasible set may not be found by standard optimization techniques. In general, multiplicative programming problem can be solved by using an outer approximation algorithm [6].

In the most literature [3–7], so called convex multiplicative programming problem which is a minimization of a product of several convex functions has been studied.

$$\min f(x) = \prod_{j=1}^{p} f_j,$$

subject to

$$g_j(x) \leq 0, \quad j = 1, 2, ..., m,$$

where $f_j : \mathbb{R}^n \to \mathbb{R}, \quad j = 1, 2, ..., p$ and $g_j : \mathbb{R}^n \to \mathbb{R}, \quad j = 1, 2, ..., m$ are convex functions.

The following result is well known [5] for solving problem (1.1)-(1.2). First, the following master problem is solved.

$$\min F(x, \lambda) = \sum_{j=1}^{p} \lambda_j f_j,$$

subject to

$$g_j(x) \leq 0, \quad j = 1, 2, ..., m,$$

$$\prod_{j=1}^{p} \lambda_j \geq 1, \quad \lambda = (\lambda_1, \lambda_2, ..., \lambda_m) \geq 0.$$

Lemma 1.1 [5] Let $(x^*, \lambda^*)$ be an optimal solution of (1.3)-(1.4). Then $x^*$ is optimal to (1.1)-(1.2).

Minimization of product of two convex functions was studied in [2]. Unlike this, in this paper we consider a problem of maximizing a product of two concave functions.
2 Quasiconcave functions

Definition 2.1 A function $f: \mathbb{D} \to \mathbb{R}$ is said to be quasiconcave on a convex set $\mathbb{D} \subset \mathbb{R}^n$ if
\[
f(\alpha x + (1 - \alpha)y) \geq \min\{f(x), f(y)\}
\]
holds for all $x, y \in \mathbb{D}$ and $\alpha \in [0, 1]$.

If $f$ is quasiconcave then $-f$ is called quasiconvex.

Theorem 2.1 [2] A function $f: \mathbb{D} \to \mathbb{R}$ is quasiconvex on $\mathbb{D}$ if and only if the set
\[
L_c(f) = \{x \in \mathbb{D} | f(x) \geq c\}
\]
is convex for all $c \in \mathbb{R}$.

Consider a problem of maximizing the product of two concave functions
\[
\max_{x \in \mathbb{D}} f = f_1 \cdot f_2 \tag{2.1}
\]
where $f_1, f_2: \mathbb{D} \to \mathbb{R}$ are positive defined concave functions on a convex set $\mathbb{D} \subset \mathbb{R}^n$.

Lemma 2.1 [2] The function $f$ is quasiconcave on $\mathbb{D} \subset \mathbb{R}^n$.

Proof. Introduce the new function $\varphi: \mathbb{D} \to \mathbb{R}$ as $\varphi = \frac{1}{f}$.

Obviously, $\varphi$ is convex and positive on $\mathbb{D}$. Then function $f$ is: $f = \frac{1}{\varphi}$.

Define the set $L_c(f)$:
\[
L_c(f) = \{x \in \mathbb{D} | f(x) \geq c\}
\]
for all positive $c \in \mathbb{R}^+$.

We show that $L_c(f)$ is convex. Take points $x, y \in L_c(f)$ and $\alpha \in [0, 1]$.

Then
\[
f_1(x) - c\varphi(x) \geq 0, \quad f_1(y) - c\varphi(y) \geq 0.
\]

Taking into account concavities of functions $f_1$ and $\varphi$, we compute
\[
f_1(\alpha x + (1 - \alpha)y) - c\varphi(\alpha x + (1 - \alpha)y).
\]

\[
f_1(\alpha x + (1 - \alpha)y) - c\varphi(\alpha x + (1 - \alpha)y) \geq \alpha f_1(x) + (1 - \alpha)f_1(y) - \alpha c\varphi(x) - (1 - \alpha)c\varphi(y) = \alpha[f_1(x) - c\varphi(x)] + (1 - \alpha)[f_1(y) - c\varphi(x)] \geq 0
\]

which is equivalent to $f(\alpha x + (1 - \alpha)y) \geq c$ and means that $\alpha x + (1 - \alpha)y \in L_c(f)$, $c > 0$.

Now we are ready to formulate global optimality conditions for problem (2.1).

On the other hand, problem (2.1) can be treated equivalently, as quasiconvex minimization problem
\[
\min_{x \in \mathbb{D}} (-f) \tag{2.2}
\]

If we apply of Theorem 2.1 [1] to problem (2.2), then we obtain global optimality conditions for the problem in the following proposition.

Theorem 2.2 Let $z$ be a global solution to problem (2.1) and $E_c(f) = \{y \in \mathbb{R}^n | f(y) = c\}$. Then
\[
(f'(x), x - y) \leq 0 \quad \text{for all} \quad y \in E_f(z)(f) \quad \text{and} \quad x \in \mathbb{D}. \tag{2.3}
\]
If, in addition, \( \lim_{\|x\| \to \infty} f(x) = +\infty \) and \( f'(x + \alpha f'(x)) \neq 0 \) holds for all \( x \in D \) and \( \alpha \geq 0 \), then condition (2.3) is sufficient for \( z \in D \) to be a global solution to problem (2.1).

**Conclusion**

We considered the problem of maximizing a product of two concave functions. The problem is nonconvex. We reduced the problem to quasiconvex minimization problem and then applied a result in [1] obtaining new global optimality conditions.

**References**


MATHEMATICAL MODELLING OF FORMING TEETH OF A
BEVEL GEAR WITH ECCENTRICALLY-CYCLOIDAL GEARING

Alexey M. Bubentchikov
Mathematics and Mechanics Department, Tomsk State University, Tomsk, Russia
michael121@mail.ru

Nikolay R. Shcherbakov
Mathematics and Mechanics Department, Tomsk State University, Tomsk, Russia
nrs@math.tsu.ru

Valentina B. Tsyrenova
Institute of Mathematics, Buryat State University, Ulan-Ude, Russia
v.ts@mail.ru

Abstract

Solving the problem of analytical description for geometry of the surface contact (the processed part and the instrument’s surface) is necessary for finding optimal methods and tools of processing parts on multicoordinate NC units. In this respect, a mathematical model for forming a bevel driving gear details with EC-gearing was constructed in this work using milling cutters of two types. Systems of equations for the determination of a cutter location at the contact point were derived. Analytical and numerical solutions of these systems were found. A complex of special computer programs to control the process of a cutter movement while processing a part was developed.

Keywords: Formbuilding, Eccentrically-cycloidal (EC) Gearing, Surface Contact

1. Introduction

Forming a part surface of a driving gear, in the most general formulation, is the task of drawing two surfaces together (the surface of an in-process part and that of a tool), which is performed in space up to the state of the first order contact, when the surfaces have a common tangent plane. A great contribution to the development of the surface formbuilding theory for
machine processing of parts was made by S.P. Radzevich [1] who used the differential-geometric method. This method is applied in this paper to solve one of formbuilding problems, namely the problem of defining the parameters of formbuilding kinematics when the machined surface of a part and the original instrumental surface of a tool are given. The problem is solved for parts of a bevel gear with eccentrically-cycloidal gearing (ec-gearing), which was developed and patented by designers of the firm «Technology market» (Tomsk, Russia) [2]. Unlike classical involute gearing, teeth profiles in EC-gearing are a cycloidal curve and a circle. The surface of the instrument is a sphere or a torus.

2. Parametric equations for the surfaces of parts

Parametric equations for the surfaces of parts

\[
x = f_1(u, v), \\
y = f_2(u, v), \\
z = f_3(u, v)
\]

will be written in the form of vector functions of two variables, i.e. as :

\[
\vec{r}(u, v) = \begin{pmatrix} f_1(u, v) \\ f_2(u, v) \\ f_3(u, v) \end{pmatrix}
\]

Let a sphere of a radius \( R \) be given, which can be defined by the vector function of two arguments, \( u \) and \( v \), taking values from 0 to \( 2\pi \):

\[
\vec{S}(u, v) = \begin{pmatrix} R \cos u \cos v \\ R \cos u \sin v \\ R \sin u \end{pmatrix}
\]

Let us denote the number of teeth of the input part (hereinafter – the tooth wheel) as \( z_1 \), the number of teeth of the output part (hereinafter - the wheels) as \( z_2 \), and the gear ratio as \( n \):

\[
n = \frac{z_2}{z_1}
\]

The cycloidal curve on the sphere is described by a point of a circle with a radius \( \varepsilon \) which lies on the sphere, with simultaneous rotation of the circle around its center at an angle \( \alpha \) and around the axis of wheel rotation (axis OZ) at an angle \( \frac{\alpha}{n} \). Such circle can be set by turning the radius vector of the point.
The wheel surface is formed by the movement of the curve in space, which is called the detail profile. In the bevel EC-gear the wheel profile is the equidistant line of the cycloidal curve on the sphere, i.e. the envelope of the circles clump with a radius $\rho<\varepsilon$ lying on the sphere, the centers of which are projected from the center of the sphere to the points of this curve (1). To construct the equidistant line it is only necessary to turn radius vectors of each point of the cycloidal curve within normal planes of this curve at an angle

$$\gamma = \arccos \frac{\sqrt{R^2 - \rho^2}}{R}.$$  

Taking orthogonal vectors $\vec{Sn}(\alpha)$ and the vector product $\vec{Sn}(\alpha) \times \vec{Sn}'(\alpha)$ as the basis vectors within the normal plane for the point of the curve (1) let us write the equidistant line of the cycloidal curve on the sphere in the form of:

$$\vec{Ev}(\alpha) = \cos \gamma \cdot \vec{Sn}(\alpha) + R \cdot \sin \gamma \frac{\vec{Sn}(\alpha) \times \vec{Sn}'(\alpha)}{\left| \vec{Sn}(\alpha) \times \vec{Sn}'(\alpha) \right|}. \quad (2)$$

To construct the wheel surface it is necessary to turn the equidistant line of the cycloidal curve around the axis $OZ$ at an angle $\frac{\nu}{n}$, while simultaneously decreasing the radius of the sphere
on which it is located from $R$ to $(R-ir)$ (where $lr$ is the width of the input detail). This leads to writing the equation of the wheel surface as a vector function of two arguments:

$$
\overline{FK}(v,\alpha) = C(v) \begin{pmatrix}
\cos \left( \frac{v}{n} \right) & -\sin \left( \frac{v}{n} \right) & 0 \\
\sin \left( \frac{v}{n} \right) & \cos \left( \frac{v}{n} \right) & 0 \\
0 & 0 & 1
\end{pmatrix} \overline{E}(\alpha), \tag{3}
$$

where

$$
C(v) = 1 - c_1v, \quad c_1 = \frac{lr \cdot z_i}{R \cdot 2\pi}
$$

and the parameter $v$ varies from 0 to $\frac{2\pi}{z_i}$.

When $v=0$ the coordinate line of the surface $\overline{FK}(v,\alpha)$ lies on the sphere of the radius $R$, and when $v = \frac{2\pi}{z_i}$ - on the sphere of the radius $\frac{R-ir}{R}$.

The tooth surface of the input parts is obtained as follows. A sphere of a radius $R$ and a circular cone with the vertex in the center of this sphere, crossing it along the circle of a radius $\epsilon$, are given. The tooth surface is formed by the circles lying on the concentric spheres of the decreasing radii, with the vertexes lying on the conical helix, moreover, the closer to the center of the sphere, the shorter the radii of the circle. The biggest working circle of the tooth is found on the sphere of the radius $R$ and it has a radius $\rho$. This circle can be defined as a vector function:

$$
\overline{okr}(\alpha) = \begin{pmatrix}
\frac{\sqrt{R^2-\epsilon^2} \cdot \sqrt{R^2-\rho^2} + \epsilon \cdot \rho \cdot \sin \alpha}{R} \\
\rho \cdot \cos \alpha \\
-\epsilon \cdot \sqrt{R^2-\rho^2} + R \cdot \epsilon \cdot \rho \cdot \sin \alpha
\end{pmatrix}. \tag{4}
$$

Then the equation of the tooth surface of a gear wheel can be written as:

$$
\overline{Fs}(v,\alpha) = C(v) \begin{pmatrix}
1 & 0 & 0 \\
0 & \cos v & -\sin v \\
0 & \sin v & \cos v
\end{pmatrix} \overline{okr}(\alpha). \tag{5}
$$

The input and output parts of a bevel gear with perpendicular concurrent axes of rotation are shown in Fig.1.
3. Formbuilding of parts with a spherical cutter

The proposed algorithm of controlling movement of a spherical cutter is suitable for cutting both input and output parts. Therefore, in the subsequent reasoning we will use the notation $\bar{F}(v, \alpha)$ considering that $\bar{F}(v, \alpha) = Fk(v, \alpha)$ (see (3)), in case of a wheel surface, and $\bar{F}(v, \alpha) = Fs(v, \alpha)$ (see (5)) – in case of a gear wheel. The cutter forms a part by serial following coordinate lines of a part surface, i.e. the surface of the cutter must at every moment of movement touch a defined coordinate line of a surface being cut. Let us denote by $Rc_i$ and $Rc_o$ ($Rc_i > Rc_o$) the distance from the center of the sphere (the cutter) to the axis of wheel rotation at the beginning and at the end of processing, moreover, these values are defined so that the diminution $Rc_i - Rc_o$ is slightly bigger than the width of parts $lr$. Thus, during the process of movement the center of the cutter is on the surface of a cylinder with a variable radius, the axis being the rotation axis of the wheel. Let us write the equations of this cylinder in the form of a vector function:

$$\overrightarrow{Cl}_k(u,v) = \begin{pmatrix} Rx_i \cos u \\ Rx_i \sin u \\ v \end{pmatrix}.$$  

Here $Rx_i$ denotes the variable radius of the cylinder:

$$Rx_i = Rc_i - \left( \frac{Rc_i - Rc_o}{K} \right) \cdot k,$$

where $K$ is the number of steps of changing the radius of the cylinder and $k$ is an integer number ranging from 0 to $K$.

For each value of this radius from $Rc_i$ to $Rc_o$ it is necessary to find the coordinates of the sphere center (that of the cutter) which touches the defined line on the working surface. The line is defined by a fixed value of the parameter $\alpha = \alpha_0$ in the equations of the surface (3) or (5). This
touch is provided by 1) the presence of a common point of a sphere with the diameter $df$ and the part surface, which can be written in vector form as:

$$\left| C_{F_i}(u,v) - F(v,a0) \right| = \frac{df}{2}$$  \hspace{1cm} (6)

and 2) parallelism of standards to these surfaces at the contact point (the common tangent plane). Since the standard to the sphere is defined by the vector $\overrightarrow{C_{F_i}(u,v) - F(v,a0)}$ and the standard to the part surface is perpendicular to the tangent lines of two coordinate lines on this surface, the condition 2) leads to two more equations:

$$\left[ C_{F_i}(u,v) - F(v,a0) \right] \cdot \frac{\partial F(v,a0)}{\partial v} = 0,$$  \hspace{1cm} (7)

$$\left[ C_{F_i}(u,v) - F(v,a0) \right] \cdot \frac{\partial F(v,a0)}{\partial u} = 0$$  \hspace{1cm} (8).

Thus we obtained the system (6) - (8) of three equations for three unknowns: $u$, $v$, $v$. Its analytical solution leads to a transcendental equation for $v$, the roots of which are numerically found using embedded software programs of the package MathCAD. With the specially developed algorithm for finding initial approximations the roots are found with high accuracy (about $10^{-12}$ mm). Fig.2 shows the position of a spherical cutter with the center lying on a defined cylinder when the cutter touches the tooth surface of the wheel at a point of the defined line on this surface.

![Figure 2. A spherical cutter touching the output part surface at a point of a defined line](image)

The figure shows the cylinder on which the center of the cutter is lying.

4. Formbuilding of parts with a torus cutter

When forming a part surface of a smaller size, it is advisable to use not a spherical cutter but, instead, a cutter which has a torus as the cutting unit, as, due to this, a smaller radius of curvature of the cutter’s cutting surface is achieved.
Let us define the torus surface, assuming that the cutter axis is the axis OZ, as a vector function of two arguments:

$$\overrightarrow{tor}(u, v) = \begin{pmatrix}
\left(\frac{df}{2} - rf\right) \cos v + rf \cos u \cos v \\
\left(\frac{df}{2} - rf\right) \sin v + rf \cos u \sin v \\
rf \sin u
\end{pmatrix}$$

(9)

where \(df\) is the cutter diameter and \(rf\) is the radius of curvature of the torus lateral surface.

When processing the surface of a wheel tooth, the cutter moves in space maintaining a parallel position of the axis towards the axis of the wheel rotation (the axis OZ). When the cutter touches the wheel surface at a point of the coordinate line defined by a fixed value of the parameter \(\alpha = \alpha_0\) in the equations of this surface (3) the position of the cutter is determined by the values of three displacements of the torus surface (9) against the coordinate axes. Defining the values of these displacements as \(Rx, Ry, Rz\) we write the equations of the cutter torus surface while touching the wheel tooth in the form of:

$$\overrightarrow{TR}(u, v) = \begin{pmatrix}
\left(\frac{df}{2} - rf\right) \cos v + rf \cos u \cos v + Rx \\
\left(\frac{df}{2} - rf\right) \sin v + rf \cos u \sin v + Ry \\
rf \sin u + Rz
\end{pmatrix}$$

(10)

Let us define the discrete set of points on the line \(\alpha = \alpha_0\) by the values of the parameter \(v\) in the equation of the wheel surface (3) with the formula:

$$v_k = \frac{2\pi k}{zK},$$

where \(K\) is the number of steps of changing the parameter \(v\) and \(k\) is an integer number ranging from 0 to \(K\). The problem of determining the location of the cutter while processing the wheel tooth is to find the values of displacements \(Rx, Ry, Rz\) for each value of \(k\), i.e. for each of the \(K\)-points on the line \(\alpha = \alpha_0\), basing on the conditions of contact between the surfaces of the cutter (10) and the part (3). These conditions are: 1) the presence of a common point of the surfaces (10) and (3) and 2) parallelism of standards to the surfaces at the contact point. From the first condition, equating the right-hand sides of (10) and (3), we obtain the expression of displacements \(Rx, Ry, Rz\) through the parameters \(u\) and \(v\):

$$\begin{pmatrix}
Rx \\
Ry \\
Rz
\end{pmatrix} = Fk(v_k, \alpha_0) - \overrightarrow{tor}(u, v).$$

(11)

To write the formulae resulting from the condition 2), let us find a standard vector to the surface of the wheel tooth (9) in a usual way:
\[ \mathbf{N}(u,v) = \frac{\partial \mathbf{t}(u,v)}{\partial u} \times \frac{\partial \mathbf{t}(u,v)}{\partial v} \begin{bmatrix} \cos u \cos v \\ \cos u \sin v \\ \sin u \end{bmatrix} \]

Similarly, differentiating the vector function (3) by the parameters \(u\) and \(v\), let us find a vector normal to the surface of the tooth wheel:

\[
\mathbf{N}(u,v) = \begin{bmatrix} f_1(u,v) \cos \left( \frac{v}{n} \right) + f_2(u,v) \sin \left( \frac{v}{n} \right) \\ f_1(u,v) \sin \left( \frac{v}{n} \right) - f_2(u,v) \cos \left( \frac{v}{n} \right) \\ f_3(u,v) \end{bmatrix},
\]

where it is indicated that:

\[
f_6(u,v) = \left( E_v(a)_i, E_v'(a)_i \right)_0 + E_v(a)_i, E_v'(a)_i \right)_i C(\nu) \begin{bmatrix} C(\nu) \\ E_v(a)_i, E_v'(a)_i \right)_i \end{bmatrix},
\]

\[
f_1(u,v) = -E_v(a)_i, E_v'(a)_i \begin{bmatrix} C(\nu) \\ E_v(a)_i, E_v'(a)_i \right)_i \end{bmatrix},
\]

\[
f_2(u,v) = -E_v(a)_i, E_v'(a)_i \begin{bmatrix} C(\nu) \\ E_v(a)_i, E_v'(a)_i \right)_i \end{bmatrix},
\]

(12)

In the formulae (12) the scalar functions \(E_v(a)_i\) and \(E_v'(a)_i\) \((i = 0,1,2)\) are the coordinates of the vector function \(\mathbf{E}(a)_i\) from (2) and its derivative with respect to the parameter \(a\), \(C(\nu)\) and \(c_i\) being introduced in (3).

From the condition 2) of standards parallelism we obtain two relations with the parameters \(u\) and \(v\) from which it is possible to express these parameters through \(u_0\) and \(v_0\). Thus, for each point \((u_0, v_0)\) on the line \(a = a_0\) we get the values of the parameters \(u\) and \(v\). If we introduce them into (10), we obtain the coordinates of the contact point for the cutter and the part, and when we introduce these values into (11), we find the values of displacements \(R_x, R_y, R_z\) which determine the cutter position in space at the moment of contact with the part. Fig.3 shows a fragment of a tooth surface of the wheel touching a torus cutter at a point of the line \(a = a_0\).
Let us write the equations of the working surface of the gear wheel tooth (5) in expanded form:

$$\overline{F} \overline{s}(v, \alpha) = C(v) \left\{ \begin{array}{c} okr(\alpha) \\ okr(\alpha), \cos v - okr(\alpha), \sin v \\ okr(\alpha), \sin v + okr(\alpha), \cos v \end{array} \right\},$$

(13)

where $okr(\alpha)_i$ $(i = 0, 1, 2)$ are the coordinates of the vector function $\overline{okr}(\alpha)$ which gives the greatest working circle of the gear wheel (4). The parameter $\alpha$ on the surface of the gear wheel is considered to be given ($\alpha = \alpha_0$), thus, on this surface we find a circular helix which is touched by the cutter while working a part. The standard vector at an arbitrary point on the surface (13) has the form:

$$\overline{Ns}(v, \alpha) = \left\{ \begin{array}{c} g_o(v, \alpha) \\ g_1(v, \alpha) \cos v + g_2(v, \alpha) \sin v \\ g_1(v, \alpha) \sin v - g_2(v, \alpha) \cos v \end{array} \right\},$$

where it is indicated that:

$$g_o(v, \alpha) = okr'(\alpha)_1 \left( okr(\alpha), C(v) - okr(\alpha), c_1 \right) + okr'(\alpha)_2 \left( okr(\alpha), C(v) + okr(\alpha), c_1 \right),$$

$$g_1(v, \alpha) = -okr'(\alpha)_2 okr(\alpha)_0 c_1 - okr'(\alpha)_3 \left( okr(\alpha), C(v) - okr(\alpha), c_1 \right),$$

$$g_2(v, \alpha) = -okr'(\alpha)_1 okr(\alpha)_0 c_1 + okr'(\alpha)_3 \left( okr(\alpha), C(v) + okr(\alpha), c_1 \right),$$

and $okr'(\alpha)_i$ $-$ $(i = 0, 1, 2)$ are the coordinates of the derivative vector function $\overline{okr}(\alpha)$.

When processing the surface of the gear wheel tooth, the cutter axis lies in a plane $\Pi$ which is perpendicular to the rotation axis of the gear wheel (the axis $OX$) and is displaced together with this plane along the axis $OX$ for a value $Rx$. Furthermore, the cutter axis rotates in the plane $\Pi$ at an
angle \( \phi \) and the torus surface is displaced along this axis for a value \( Ht \). These three values determine the position of the cutter at each moment of processing a part (i.e. for each point \( \overrightarrow{FS}(v_k, a0) \) of the determined line for \( \alpha = \alpha_0 \) on the surface being processed) and they must be found from the condition of contact between the interacting surfaces. Let us write the equations of the cutter surface with an account of these displacements in the form:

\[
\overrightarrow{Tr}(u, v, \phi, Rx, Ht) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \phi & \sin \phi \\ 0 & -\sin \phi & \cos \phi \end{pmatrix} \begin{pmatrix} 0 \\ \overrightarrow{tor}(u, v) + \left( \begin{array}{c} 0 \\ \alpha_0 \\ Ht \end{array} \right) \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ Rx \end{pmatrix}
\]

Then the standard unit vector of the torus surface at the point with curvilinear coordinates \((u, v)\) has the form:

\[
\overrightarrow{Nt}(u, v, \phi) = \begin{pmatrix} \cos u \cos \nu \\ \sin \phi \sin u + \cos \phi \cos u \sin \nu \\ \cos \phi \sin u - \sin \phi \cos u \sin \nu \end{pmatrix}
\]

To find the contact points of the part surface and the torus (the cutter) we use two conditions: parallelism of standards of these surfaces at the contact point and this point belonging to both surfaces.

From the condition of the standards parallelism:

\[
\overrightarrow{Ns}(v_k, a0) \times \overrightarrow{Nt}(u, v, \phi) = 0
\]

at \( \alpha = \alpha_0 \) we obtain two equations with three unknowns \( u, v, \phi \), from which we can express \( u \) and \( v \) through \( \phi \):

\[
\tan(u) = \frac{\sin(v_k + \phi)g_z(v_k, a0) - \cos(v_k + \phi)g_y(v_k, a0)}{\sqrt{g_y(v_k, a0)^2 + g_z(v_k, a0)^2}}, \quad \tan(v) = \frac{\cos(v_k + \phi)g_z(v_k, a0) + \sin(v_k + \phi)g_y(v_k, a0)}{g_y(v_k, a0)}
\]

(14)

The condition of having a common point of the surfaces:

\[
\overrightarrow{Tr}(u, v, \phi, Rx, Ht) = \overrightarrow{FS}(v_k, a0)
\]

leads to three more equations with three unknowns \( \phi, Rx \) and \( Ht \) (after introducing formulae (14) into these equations). An analytical solution of this system leads to a transcendental equation for \( \phi \), the roots of which are numerically found using embedded programs of the software package MathCAD. With the specially developed algorithm for finding initial approximations the roots are found with high accuracy (about \( 10^{-12} \) mm). Fig.4 shows the position of a torus cutter touching the surface of a gear wheel tooth at a point of a determined line on this surface.
5. Conclusion

The problems solved in this paper were used in development of computer programs for CNC machines that control movements of milling cutters in formbuilding of parts of different mechanisms with EC-engagement. These programs are successfully used by the firm «Technology market» (Tomsk, Russia) in the manufacture of traction gear boxes prototypes for rail transport and a two-stage gear [3], which was successfully tested at the company SEW (Germany).

References


List of Figures
1 The input and output parts of a bevel gear in EC-gearing ........................................... 4

2 A spherical cutter touching the output part surface at a point of a defined line.......................... 5

3 Fragment of a tooth surface of the wheel touching a torus cutter ...........................................7

4. Contact interaction of the surface of a gear wheel tooth with a torus cutter .................................8
THE FIXED POINT METHOD IN POLYNOMIAL WITH RESPECT TO STATE PROBLEM OF PARAMETRIC OPTIMIZATION

©A.S. Buldaev

Russia, Ulan-Ude, Buryat State University

G. Ankhbayar

Mongolia, Ulaanbaatar University

I.-Kh.D. Khishektueva

Russia, Ulan-Ude, Buryat State University

Abstract. A new approach for improving control parameters of polynomial with respect to state systems is proposed. The approach is based on the construction and solution of the problem of fixed point defined by the projection operator. The procedure can improve controls satisfying maximum principle.

Keywords: parametric optimization, fixed point problem

1 Polynomial with respect to state problem of parametric optimization

We consider the problem of optimizing a dynamic system by control parameters

\[ \Phi(u) = \varphi(x(t_1)) + \int_T F(x(t), u, t) dt \rightarrow \min_{u \in U}, \]  \hspace{1cm} (1)

\[ \dot{x}(t) = f(x(t), u, t), \ x(t_0) = x^0, \ u \in U, \ t \in T = [t_0, t_1], \]  \hspace{1cm} (2)

\[ u \rightarrow \min \]  \hspace{1cm} (3)

\[ \dot{p}(t) = (H(x(t), u, t) - \dot{x}(t), p(t)), \ p(t_0) = 0, \ t \in T. \]  \hspace{1cm} (4)

\[ u \rightarrow \min \]  \hspace{1cm} (5)

\[ \dot{w}(t) = \psi(f(x(t), u, t) - \dot{x}(t)), \ w(t_0) = 0, \ t \in T. \]  \hspace{1cm} (6)

\[ u \rightarrow \min \]  \hspace{1cm} (7)

\[ \dot{y}(t) = \varphi(x(t)), \ y(t_0) = 0, \ t \in T. \]  \hspace{1cm} (8)

\[ u \rightarrow \min \]  \hspace{1cm} (9)

\[ \dot{z}(t) = \dot{y}(t) + \psi(f(x(t), u, t) - \dot{x}(t)), \ z(t_0) = 0, \ t \in T. \]  \hspace{1cm} (10)

\[ u \rightarrow \min \]  \hspace{1cm} (11)

\[ \dot{w}(t) = \psi(f(x(t), u, t) - \dot{x}(t)), \ w(t_0) = 0, \ t \in T. \]  \hspace{1cm} (12)

\[ u \rightarrow \min \]  \hspace{1cm} (13)

\[ \dot{y}(t) = \varphi(x(t)), \ y(t_0) = 0, \ t \in T. \]  \hspace{1cm} (14)

\[ u \rightarrow \min \]  \hspace{1cm} (15)

\[ \dot{z}(t) = \dot{y}(t) + \psi(f(x(t), u, t) - \dot{x}(t)), \ z(t_0) = 0, \ t \in T. \]  \hspace{1cm} (16)

\[ u \rightarrow \min \]  \hspace{1cm} (17)

\[ \dot{w}(t) = \psi(f(x(t), u, t) - \dot{x}(t)), \ w(t_0) = 0, \ t \in T. \]  \hspace{1cm} (18)

\[ u \rightarrow \min \]  \hspace{1cm} (19)

\[ \dot{y}(t) = \varphi(x(t)), \ y(t_0) = 0, \ t \in T. \]  \hspace{1cm} (20)

\[ u \rightarrow \min \]  \hspace{1cm} (21)

\[ \dot{z}(t) = \dot{y}(t) + \psi(f(x(t), u, t) - \dot{x}(t)), \ z(t_0) = 0, \ t \in T. \]  \hspace{1cm} (22)

\[ u \rightarrow \min \]  \hspace{1cm} (23)

\[ \dot{w}(t) = \psi(f(x(t), u, t) - \dot{x}(t)), \ w(t_0) = 0, \ t \in T. \]  \hspace{1cm} (24)

\[ u \rightarrow \min \]  \hspace{1cm} (25)
where \( u = (u_1, \ldots, u_m) \) - vector of control parameters (control) with values in a compact set \( U \subseteq \mathbb{R}^m \). Functions \( f(x, u, t), F(x, u, t) \) are polynomials of an entire degree \( k \geq 1 \) with respect to \( x \) with coefficients, depending continuously on \( u, t \), on the set \( \mathbb{R}^n \times U \times T \), function \( \varphi(x) \) - is a polynomial of degree \( k \) on \( \mathbb{R}^n \). The initial state \( x^0 \) and the interval \( T \) are fixed.

Let us form the Pontryagin function

\[
H(\psi, x, u, t) = \langle \psi, f(x, w, t) \rangle - F(x, w, t).
\]

Consider a vector standard adjoint system

\[
\dot{\psi}(t) = -H_x(\psi(t), x(t), w, t), \ t \in T
\]  \hspace{1cm} (3)

and introduce a vector modified adjoint system

\[
\dot{p}(t) = -H_x(p(t), x(t), w, t) - \frac{1}{2!} \langle H_x(p(t), x(t), w, t), z \rangle_x - \\
\ldots - \frac{1}{k!} \langle \ldots \langle H_x(p(t), x(t), w, t), z \rangle_x, z \rangle_x \ldots, z \rangle_x.
\]

\hspace{1cm} (4)

For admissible controls \( u, v \) let us denote \( \Delta_v \Phi(u) \) as an increment of the objective function; \( x(t, v), t \in T \) - is a solution of system (2) at \( w = v \), \( x(t_0, v) = x^0 \); \( \psi(t, v), t \in T \) - is a solution of system (3) at \( w = v \), \( x = x(t, v) \), \( \psi(t_1, v) = -\varphi_x(x(t_1, v)) \); \( p(t, u, v), t \in T \) - is a solution of system (4) at \( w = u \), \( z = x(t, v) - x(t, u) \) \( x(t, v) \) \( x(t, u) \) with a boundary condition

\[
p(t_1, u, v) = -\varphi_x(x(t_1, u)) - \frac{1}{2!} \langle \varphi_x(x(t_1, u)), y \rangle_x - \\
\ldots - \frac{1}{k!} \langle \ldots \langle \varphi_x(x(t_1, u)), y \rangle_x, y \rangle_x \ldots, y \rangle_x,
\]

where partial derivatives with respect to \( x \) are calculated at \( y = x(t_1, v) - x(t_1, u) \). It is evident that \( p(t, u, u) = \psi(t, u), t \in T \).
Let us set the improvement control problem for $u \in U$: to find a control $v \in U$ satisfying the condition $\Delta_v \Phi(u) \leq 0$.

Control vector of parameters in the problem (1), (2) is considered as a constant vector function of time on the interval $T$. Then from the exact formulas of increment of objective functional in the optimal control problem [1] as an obvious corollary we obtain the first and the second exact formulas of increment of objective function accordingly in the problem (1), (2).

\begin{align*}
\Delta_v \Phi(u) &= -\int_T \Delta_v H(p(t, u, v), x(t, v), u, t) dt, \quad (5) \\
\Delta_v \Phi(u) &= -\int_T \Delta_v H(p(t, v, u), x(t, u), u, t) dt. \quad (6)
\end{align*}

Let us introduce a mapping using the formula

\[ W^*(u, v) = \arg\max_{w \in U} \int_T H(p(t, u, v), x(t, v), w, t) dt, \quad u \in U, \ v \in U. \quad (7) \]

According to the formula (5) for optimality of control $u \in U$ it is sufficient (and necessary) that

\[ \int_T \Delta_v H(p(t, u, v), x(t, v), u, t) dt \leq 0, \ v \in U. \]

To fulfill the last inequality it is sufficient to require a fulfillment of the condition

\[ u = W^*(u, v), \ v \in U \quad (8) \]

The sufficient optimality condition, based on the formula (6), accordingly has the form

\[ u = W^*(v, u), \ v \in U \quad (9) \]

In case of differentiability of the function $H$ with respect to $u$ and convexity of the set $U$ the necessary optimality condition for the control $u \in U$ in the problem (1), (2) has the form of known differential maximum principle (DPM) [2]

\[ u = \arg\max_{w \in U} \int_T \langle H_u(\psi(t, u), x(t, u), u, t), w \rangle dt. \quad (10) \]
Let us extract the condition

\[ u = W^*(u, u), \]  

(11)

which is obtained from sufficient conditions (8), (9) at \( v = u \). It is easy to show that (10) is a corollary of (11).

Consider an important for applications subclass of linear with respect to control problems

\[ \Phi(u) = \varphi(x(t_1)) + \int_T \langle (a(x(t), t), u) + d(x(t), t) \rangle dt \rightarrow \min_{u \in U}, \]  

(12)

\[ \dot{x}(t) = A(x(t), t)u + b(x(t), t), \quad x(t_0) = x^0, \quad u \in U, \quad t \in T, \]  

(13)

where the matrix function \( A(x, t) \), vector functions \( a(x, t), b(x, t) \), functions \( \varphi(x) \), \( d(x, t) \) are polynomial with respect to \( x \) and continuos with respect to \( t \) on the set \( R^n \times T \). The set \( U \in R^m \) is convex and compact.

In problem (12), (13) the differential maximum principle for control \( u \in U \) is represented as (11). The Pontryagin function, the exact formulas of the increment of objective function (5), (6) and the mapping \( W^*(u, v) \) have the form

\[ H(\psi, x, u, t) = H_0(\psi, x, t) + \langle H_1(\psi, x, t), u \rangle, \]

\[ H_0(\psi, x, t) = \langle \psi, b(x, t) \rangle - d(x, t), \quad H_1(\psi, x, t) = A^T(x, t)\psi - a(x, t), \]

\[ \Delta_v \Phi(u) = -\langle \int_T H_1(p(t, u, v), x(t, v), t)dt, v - u \rangle, \]

\[ \Delta_v \Phi(u) = -\langle \int_T H_1(p(t, v, u), x(t, u), t)dt, v - u \rangle, \]

\[ W^*(u, v) = \arg \max_{w \in U} \int_T H_1(p(t, u, v), x(t, v), t)dt, w \rangle, \quad u \in U, \quad v \in U. \]

For admissible control \( u \in U \) let us define the projection \( u^\alpha, \alpha > 0 \) using the relation

\[ u^\alpha = P_U(u + \alpha \int_T H_1(\psi(t, u), x(t, u), t)dt), \]

41
where $P_U$ - is a projection operator on the set $U$ in the Euclidean norm.

Differential maximum principle in problem (12), (13) for control $u \in U$ on the basis of projection property (16) is represented in the form

$$u = u^\alpha, \alpha > 0.$$  \hspace{1cm} (14)

Note, that to fulfill (11) it is sufficient to verify the condition (14) at least for one $\alpha > 0$. Conversely, from the condition (11) it follows that (14) is fulfilled for all $\alpha > 0$.

2 Fixed point methods in parametric optimization

The exact formulas (5), (6) enable us to construct methods of fixed point in problem (1), (2) using the operation for the maximum (7).

Let us consider the improvement control problem for $u^0 \in U$.

The first fixed point method: for given $u^0 \in U$ we define the mapping $W^*_1$ using the relation $W^*_1(v) = W^*(u^0, v), v \in U$ and find a solution $v = v(u^0)$ of a fixed point problem

$$v = W^*_1(v).$$  \hspace{1cm} (15)

As $W^*_1(v) \in U$, then the obtained control $v = v(u^0)$ is admissible ($v \in U$). By virtue of definition of the mapping $W^*_1$ we obtain

$$\int_T H(p(t, u^0, v), x(t, v), v, t)dt \geq \int_T H(p(t, u^0, v), x(t, v), u^0, t)dt.$$  

From here and from the formula (5) follows $\Delta_v \Phi(u^0) \leq 0$.

The second fixed point method: for given $u^0 \in U$ define the mapping $W^*_2$ using the relation $W^*_2(v) = W^*(v, u^0), v \in U$ and find a solution of a fixed point problem.

$$v = W^*_2(v).$$  \hspace{1cm} (16)
By analogy with the first procedure (15), we find that the solution of equation (16) in the second method is admissible and provides an improvement.

Thus, the fixed point methods consist in search for fixed points of corresponding mappings $W_1^*$ and $W_2^*$.

Let us consider the sets of fixed points $V_1(u^0) = \{v \in U : v = W_1^*(v)\}$ and $V_2(u^0) = \{v \in U : v = W_2^*(v)\}$ in corresponding procedures of improvement. If $u^0 \in V_1(u^0)$, then $u^0$ satisfies condition (11). Conversely, if $u^0$ satisfies condition (11), then $u^0$ is a solution of equation (15), that is $u^0 \in V_1(u^0)$. Hence, the condition (11) for $u^0 \in U$ characterizes inclusion $u^0 \in V_1(u^0)$. Similarly the condition (11) is equivalent to inclusion $u^0 \in V_2(u^0)$. Thus, the following statement is true.

**Lemma 1.** Control $u^0 \in U$ satisfies the condition (11) then and only then, when $u^0 \in V_1(u^0)$ (respectively $u^0 \in V_2(u^0)$).

**Corollary 1.** In problem (12), (13) control $u^0 \in U$ satisfies the differential maximum principle then and only then, when $u^0 \in V_1(u^0)$ (respectively $u^0 \in V_2(u^0)$).

Thus, the absence of fixed points in methods indicates a non-optimality of control $u^0 \in U$ in problem (12), (13).

Property of non-uniqueness of fixed points in methods is a necessary condition for strict improvement of controls, satisfying the DMP in the problem (12), (13).

Indicate the conditions, under which there is a strict improvement $\Phi(v) - \Phi(u^0)$ at $v \in V_1(u^0)$ ($v \in V_2(u^0)$). Restrict ourselves to the case of a linear with respect to control problem (12), (13).

Define the vector-functions in the first and second methods respectively in the form

$$g_1(v) = \int_T H_1(p(t, u^0, v), x(t, v), t) dt, \ v \in U,$$
\[ g_2(v) = \int_T H_1(p(t, v, u^0), x(t, u^0), t) dt, \quad v \in U. \]

Herewith mappings \( W_1^* \) and \( W_2^* \) in methods of fixed points take respectively the form

\[ W_1^*(v) = \arg\max_{w \in U} \langle g_1(v), w \rangle, \quad v \in U, \]

\[ W_2^*(v) = \arg\max_{w \in U} \langle g_2(v), w \rangle, \quad v \in U. \]

The first (14) and the second (15) formulas of the increment of objective function at controls \( u^0 \in U, \ v \in U \) are written in the form

\[ \Delta_v \Phi(u^0) = -\langle g_1(v), v - u^0 \rangle, \]

\[ \Delta_v \Phi(u^0) = -\langle g_2(v), v - u^0 \rangle. \]

Assume that \( v \in U \) - is a fixed point in method. Then from the formulas of increment follows that the strict improvement is guaranteed (including for control \( u^0 \), satisfying DMP), if vectors \( v - u^0 \) and \( g_1(v) \) (respectively \( g_2(v) \)) are not orthogonal. In particular, for scalar control \( (m = 1) \) strict improvement is guaranteed, if \( v \) is not equal to \( u^0 \) and function \( g_1 \) (respectively \( g_2 \)) is not equal to zero at the point \( v \).

The points \( v \in U \), at which function \( g_1(g_2) \) is equal to zero, obviously are fixed points of mapping \( W_1^*(W_2^*) \). Let us call these fixed points singular, the others - nonsingular. The condition of strict improvement in problem (12), (13) in case of scalar control can be formulated as follows: if \( v \in U \) - is a nonsingular fixed point in procedure of improvement, different from \( u^0 \), then \( \Phi(v) < \Phi(u^0) \).

Note, that in the problem (12), (13) nonsingular fixed points can only be the boundary points of the set \( U \).
Corollary 2. Singular fixed points in the methods of fixed point do not give strict improvement of control in linear with respect to control problem (12), (13).

The proposed methods indicate the possibility in principle of realization of nonlocal improvement of control parameters in considered class of polynomial with respect to state problems. The complexity of construction of vector of control parameters that improves is determined by the complexity of solving the problem of fixed point of auxiliary vector-functions in methods. The mapping (7) is generally determined by many-valued and discontinuous vector functions $W^*_1$ and $W^*_2$, values of which are calculated through the operation of integration of phase and adjoint systems and the operation for the maximum. Compensation for the search of fixed points of the mapping is a property of nonlocality of improvement and a possibility of strict improvement of controls that satisfy differential maximum principle.

3 Examples

Let us illustrate the work of the proposed fixed point methods on simple examples.

Example 1 (improvement of control).

$$\Phi(u) = \frac{1}{2} \int_0^2 x^2(t) dt \to \min,$$

$$\dot{x}(t) = u, \ x(0) = 1, \ u \in U = [-1, 1], \ t \in T = [0, 2].$$

In this case $H = pu - \frac{1}{2}x^2$, the modified adjoint system $\dot{p}(t) = x(t) + \frac{1}{2}y(t), \ p(2) = 0$.

Consider control $u^0 = 0$ with a corresponding phase trajectory $x(t, u^0) = 1, \ t \in T$ and a value of functional $\Phi(u^0) = 1$. Let us set a problem of improvement of control $u^0$. 

45
Apply the first fixed point method. We have

\[ x(t, v) = vt + 1, \quad p(t, u^0, v) = t + \frac{vt^2}{4} - (v + 2), \quad t \in T, \]

\[ W_1^*(v) = \text{sign}g_1(v), \quad g_1(v) = \int_0^2 p(t, u^0, v)dt = -2 - \frac{4}{3}v, \quad v \in U. \]

A unique zero of the switching function \( g_1 \) is the point \( v = -\frac{3}{2} \notin [-1, 1] \). Consequently, there are no singular fixed points. Using the method of substitution a unique nonsingular fixed point \( v = -1 \) is defined.

Let us discuss the procedure. Since \( u^0 = 0 \) does not satisfy the differential maximum principle. Since \( v = -1 \neq u^0 \) and \( g_1(-1) \neq 0 \), then \( v = -1 \) strictly improves \( u^0 = 0 \) with a value of the increment \( \Delta \Phi(u^0) = -(v - u^0)g_1(v) = -\frac{2}{3} \).

**Example 2** (improvement of control, satisfying the maximum principle).

\[ \Phi(u) = \frac{1}{2} \int_0^1 ux^2(t)dt \rightarrow \min_{u \in U}, \]

\[ \dot{x}(t) = u, \quad x(0) = 0, \quad u \in U = [-1, 1], \quad t \in T = [0, 1]. \]

Let us set a problem of improvement of control \( u^0 = 0 \) with a corresponding phase trajectory \( x(t, u^0) = 0, \quad t \in T \) and a value of the objective function \( \Phi(u^0) = 0 \).

Pontryagin function is \( H = \psi u - \frac{1}{2}ux^2 \), standard adjoint system is \( \dot{\psi}(t) = ux(t), \quad \psi(1) = 0 \). We have \( \psi(t, u^0) = 0, \quad t \in T \). From this we obtain \( H_u(\psi(t, u^0), x(t, u^0), u^0, t) = 0, \quad t \in T \). Thus, \( u^0 \) is a special control, i.e. it satisfies the differential maximum principle with degeneracy.

Apply the first method of fixed point. Define a modified adjoint system \( \dot{p}(t) = ux(t) + \frac{1}{2}uy(t), \quad p(1) = 0 \); the mapping \( W_1^*(v) = \text{sign} \int_0^1 (p(t, u^0, v) - \frac{1}{2}x^2(t, v))dt, \quad v \in U \). We obtain \( p(t, u^0, v) = 0, \quad x(t, v) = vt, \quad t \in T, \quad W_1^*(v) = \text{sign}(-\frac{1}{6}v^2), \quad v \in U. \)
A unique zero of the switching function \( g_1(v) = -\frac{1}{6}v^2 \) is the point \( v = 0 \in U \), which is the singular fixed point of the mapping \( W_1^* \). Nonsingular fixed point \( v = -1 \) strictly improves \( u^0 = 0 \) with a value of the increment \( \Delta v \Phi(u^0) = -(v - u^0)g_1(v) = -\frac{1}{6} \).

This example demonstrates the following properties of constructed methods:

- control, satisfying condition (11), is a fixed point of mapping in the methods;
- possibility of nonunique solution of the problem of fixed point in the methods;
- possibility of strict improvement of control that satisfies the differential maximum principle.

4 Conclusion

Let us distinguish the following properties of methods of fixed points that favourably distinguish them from the gradient methods.

1. Nonlocality of improvement of control parameters without procedure of variation on small parameter.

2. Possibility of strict improvement of control parameters that satisfy differential maximum principle.

In linear control problem such a possibility is conditioned by non-uniqueness of solution of the problem of fixed point in methods.

3. Possibility of improvement of control parameters on nonconvex sets in terms of absence of the property of differentiability on control for functions \( f(x, u, t), \ F(x, u, t) \).

Gradient methods do not provide such a possibility.
In the case where in procedures of fixed point there are no admissible fixed points the method does not work and we have to turn to other methods of improvement. In problem (12), (13) this means that the improvable vector of control parameters does not satisfy differential maximim principle and hence, is not optimal.

**Literature**


