

Ludvík Prouza

Annäherung nichtnegativer Funktionen mit Hilfe der Fejérschen Polynome als  
Problem der quadratischen Programmierung

*Kybernetika*, Vol. 5 (1969), No. 6, (457)--473

Persistent URL: <http://dml.cz/dmlcz/124141>

## Terms of use:

© Institute of Information Theory and Automation AS CR, 1969

Institute of Mathematics of the Academy of Sciences of the Czech Republic provides access to digitized documents strictly for personal use. Each copy of any part of this document must contain these *Terms of use*.



This paper has been digitized, optimized for electronic delivery and stamped with digital signature within the project *DML-CZ: The Czech Digital Mathematics Library*  
<http://project.dml.cz>

# Annäherung nichtnegativer Funktionen mit Hilfe der Fejérschen Polynome als Problem der quadratischen Programmierung

LUDVÍK PROUZA

Es wird eine Methode beschrieben, die — besonders mit Anwendung eines Digitalrechners — ermöglicht, nichtnegative integrierbare gerade Funktionen im Intervall  $\langle -\pi, \pi \rangle$  mit Hilfe der Fejérschen trigonometrischen Polynome im Sinne der kleinsten Quadrate zu approximieren.

## 1. EINLEITUNG

Es sei

$$(1) \quad f_n(t) = \mu_0 + \mu_1 \cos t + \dots + \mu_n \cos nt \geq 0$$

ein Fejérsches trigonometrisches Kosinuspolynom mit reellen Koeffizienten. Wie bekannt [2], existiert ein Polynom  $g_n(z)$  der komplexen Veränderlichen  $z$ , das  $n$ -ten Grades ist, reelle Koeffizienten besitzt, und es gilt

$$(2) \quad f_n(t) = |g_n(e^{it})|^2.$$

Man kann also  $g_n(z)$  (mit eventueller Verzögerung  $z^{-k}$ ,  $k > 0$ , der physikalischen Realisierbarkeit wegen) als Übertragungsfunktion eines linearen diskreten Filters betrachten (ohne Rückkopplung und mit „beschränktem Gedächtnis“).  $\sqrt{f_n(t)}$  stellt die Amplitudencharakteristik dieses Filters dar.

Es sei  $h(t)$  eine nichtnegative integrierbare gerade Funktion im Intervall  $\langle -\pi, \pi \rangle$ . Versuchen wir  $h(t)$  mit Hilfe von  $f_n(t)$  zu approximieren so, dass (bei gegebenem  $n$ )

$$(3) \quad \int_{-\pi}^{\pi} [f_n(t) - h(t)]^2 dt = \min.$$

Man kann sich unschwer überzeugen, dass man bei „direkter“ Lösung mit Hilfe der Lagrangeschen Koeffizienten auf grosse Schwierigkeiten stösst. Die Neben-

bedingungen, die entweder im Zusammenhang der Koeffizienten von  $f_n(t)$  und  $g_n(z)$  bestehen oder in Beschränkungen der Lage der Wurzeln von  $g_n(z)$ , sind schwer zu beherrschen.

In diesem Artikel wollen wir eine indirekte Lösung der Aufgabe zeigen. Wir approximieren  $h(t)$  zuerst mit einem Abschnitt der Fourierreihe und dann diesen mit einem trigonometrischen Fejérschen Polynom von gleicher Gliederanzahl.

## 2. EIN ANNÄHERUNGSSATZ

Bei Annäherung einer nichtnegativen integrierbaren geraden Funktion  $h(t)$  mit Hilfe der Fourierreihe kann der Fall vorkommen (in Bezug auf die Gestalt von  $h(t)$ ), dass alle Abschnitte der Reihe im Bereiche unseren Interesses nichtnegativ sind. Dann ist jeder Abschnitt zugleich das zugehörige Fejérsche Polynom und die Aufgabe ist trivialerweise gelöst.

Es kann jedoch auch der Fall vorkommen, dass die für uns interessanten Abschnitte irgendwo im Intervall  $\langle -\pi, \pi \rangle$  negativ sind. Mit diesem nichttrivialen Fall werden wir uns weiter beschäftigen.

Es sei nun  $\{\varphi_j(t)\}$  ( $j = 0, 1, 2, \dots$ ) ein System von orthonormierten Funktionen (in einem Intervall der reellen Veränderlichen  $t$ ). Es sei  $h(t)$  eine integrierbare Funktion in demselben Intervall. Konstruieren wir zu dieser den Abschnitt der Fourierreihe ( $n$  gegeben)

$$(4) \quad T_n(t) = \sum_{j=0}^n \lambda_j \varphi_j(t),$$

wo  $\lambda_j$  ( $j = 0, 1, 2, \dots$ ) die Fourierkoeffizienten zu  $h(t)$  sind.

Betrachten wir jetzt alle Funktionen der Gestalt

$$(5) \quad f_n(t) = \sum_{j=0}^n \mu_j \varphi_j(t),$$

welche eine gewisse Nebenbedingung erfüllen (in unserem Fall der Fejérschen trigonometrischen Polynome ist es die Beschränkung  $\sum \mu_j \varphi_j(t) \geq 0$ ).

Betrachten wir den  $n + 1$ -dimensionalen euklidischen Raum. In diesem ist  $\{\lambda_0, \lambda_1, \dots, \lambda_n\}$  ein durch  $h(t)$  bestimmter Punkt. Die Punkte, die die gegebene Nebenbedingung erfüllen, bilden im Raum eine Untermenge  $M$ .

**Satz 1.** *Es liege  $\{\lambda_0, \lambda_1, \dots, \lambda_n\}$  ausserhalb der Menge  $M$ . Enthält  $M$  einen Punkt  $\{\mu_0, \mu_1, \dots, \mu_n\}$ , dessen Entfernung vom Punkt  $\{\lambda_0, \lambda_1, \dots, \lambda_n\}$  am kleinsten ist, so approximiert die zugehörige Funktion (5) die Funktion  $h(t)$  am besten im Sinne (3). Existiert umgekehrt ein einziger Punkt  $\{\mu_0, \mu_1, \dots, \mu_n\}$  in  $M$ , in dem (3) gilt, so liegt dieser dem Punkt  $\{\lambda_0, \lambda_1, \dots, \lambda_n\}$  am nächsten.*

Beweis: Laut Voraussetzung gilt im gegebenen Intervall

$$(6) \quad \int \varphi_j^2(t) dt = 1 \quad (j = 0, 1, 2, \dots),$$

$$\int \varphi_j(t) \varphi_k(t) dt = 0 \quad (j \neq k; j, k = 0, 1, 2, \dots).$$

Weiter ist

$$(7) \quad \lambda_j = \int h(t) \varphi_j(t) dt$$

und es gilt

$$(8) \quad \int \left[ \sum_{j=0}^n \lambda_j \varphi_j(t) - h(t) \right]^2 dt = \min.$$

In Bezug auf (6) gilt für die Punkte von  $\mathcal{M}$

$$(9) \quad \int \left[ \sum_{j=0}^n \mu_j \varphi_j(t) - \sum_{j=0}^n \lambda_j \varphi_j(t) \right]^2 dt = \sum_{j=0}^n (\mu_j - \lambda_j)^2 \int \varphi_j^2 dt = \sum_{j=0}^n (\mu_j - \lambda_j)^2.$$

Weiter bekommt man mit (6), (7), (8), (9),

$$(10) \quad \int \left[ \sum \mu_j \varphi_j(t) - h(t) \right]^2 dt = \int \left\{ \left[ \sum (\mu_j - \lambda_j) \varphi_j(t) \right] + \left[ \sum \lambda_j \varphi_j(t) - h(t) \right] \right\}^2 dt =$$

$$= \sum (\mu_j - \lambda_j)^2 + \int \left[ \sum \lambda_j \varphi_j(t) - h(t) \right]^2 dt.$$

Die linke Seite in (10) ist grösser als das absolute Minimum (8) gerade um das erste Glied der rechten Seite. Dieses ist aber minimal für den Punkt aus  $\mathcal{M}$ , dessen Entfernung von  $\{\lambda_0, \lambda_1, \dots, \lambda_n\}$  die kleinste ist. Ist umgekehrt die linke Seite in (10) für ein  $\{\mu_0, \mu_1, \dots, \mu_n\}$  minimal, so ist offenbar auch das erste Glied rechts minimal. Damit ist der Satz 1 bewiesen.

### 3. EXISTENZ UND EINDEUTIGKEIT DER ANNÄHERUNG MIT FEJÉRSCHEN TRIGONOMETRISCHEN POLYNOMEN

Betrachten wir wieder die Fejérschen Kosinuspolynome

$$(11) \quad f_n(t) = \mu_0 + \mu_1 \cos t + \dots + \mu_n \cos nt \geq 0,$$

$$(t \in \langle 0, \pi \rangle).$$

Sind  $\mu_0, \mu_1, \dots, \mu_n$  die Koordinaten eines Punktes im  $n + 1$ -dimensionalen Raum,

dann stellt (11) für jedes  $t$  einen  $n + 1$ -dimensionalen Halbraum dar, der „über“ dem  $n$ -dimensionalen linearen Unterraum (diesen eingerechnet)

$$(12) \quad \mu_0 + \mu_1 \cos t + \dots + \mu_n \cos nt = 0$$

liegt. Die Menge  $M$  der Punkte  $\{\mu_0, \mu_1, \dots, \mu_n\}$ , die (11) für jedes  $t \in \langle 0, \pi \rangle$  erfüllen, ist Durchschnitt dieser Halbräume.

Es sei  $h(t)$  eine nichtnegative gerade Funktion und  $\{\lambda_0, \lambda_1, \dots, \lambda_n\}$  der ihr zugehörige Punkt, dessen Koordinaten die Fourierkoeffizienten von  $h(t)$  sind. Nach Voraussetzung verletzt dieser Punkt (11) für mindestens ein  $t \in \langle 0, \pi \rangle$  (bei Substitution  $\mu_0 = \lambda_0, \mu_1 = \lambda_1, \dots, \mu_n = \lambda_n$ ). Existiert dann ein Punkt in  $M$ , für den (3) gilt, so gilt nach dem Satz 1 für diesen Punkt  $\{\mu_0, \mu_1, \dots, \mu_n\}$

$$(13) \quad (\mu_0 - \lambda_0)^2 + \dots + (\mu_n - \lambda_n)^2 = \min$$

mit den Nebenbedingungen (11).

Man sieht, es handelt sich hier um eine Aufgabe der quadratischen Programmierung, „separabil“ nach den Veränderlichen [1], die nichtnegativ nicht zu sein brauchen. Dagegen hat die Menge der Nebenbedingungen (11) die Mächtigkeit des Kontinuums.

Bemerken wir noch, dass der Punkt  $\{\lambda_0, \lambda_1, \dots, \lambda_n\}$  nicht im ganzen Raum liegen kann; seine Koordinaten erfüllen (s. [2], S. 71) die Beschränkungen ( $\lambda_0 > 0$ )

$$(14) \quad -2 \leq \frac{\lambda_j}{\lambda_0} \leq 2 \quad (j = 1, 2, \dots).$$

Die Existenz und Eindeutigkeit der Lösung unserer Aufgabe folgt aus den nachstehenden drei Sätzen.

**Satz 2.** Eine positiv definite quadratische Form ist eine streng konvexe Funktion.

Beweis [1], S. 99.

**Satz 3.** Die Menge  $M$  ist abgeschlossen und konvex.

Beweis. Jeder der Halbräume (11) ist abgeschlossen und konvex. Also ist bekanntlich ([1], S. 51) auch ihr Durchschnitt, die Menge  $M$ , abgeschlossen und konvex. Somit ist der Satz 3 bewiesen.

**Satz 4.** Eine streng konvexe Funktion erreicht in einer konvexen abgeschlossenen Menge ein einziges (globales) Minimum in einem einzigen Punkt.

Beweis [1], S. 105–106.

Es ist klar, dass für  $\{\lambda_0, \lambda_1, \dots, \lambda_n\}$  ausserhalb  $M$  der Punkt  $\{\mu_0, \mu_1, \dots, \mu_n\}$  des Minimums auf der Grenze von  $M$  liegt. Würde er im Inneren von  $M$  liegen, so gäbe es auf der Verbindungsgeraden der Punkte  $\{\lambda_0, \dots, \lambda_n\}, \{\mu_0, \dots, \mu_n\}$  einen

anderen, der auch im Inneren von  $M$  liegen würde und für welchen (13) kleiner wäre. Das ist aber unmöglich.

#### 4. LÖSUNG DER APPROXIMIERUNGSAUFGABE ALS EINER DER QUADRATISCHEN PROGRAMMIERUNG

Eine Lösung, die sich unmittelbar darbietet, besteht darin, dass man sich statt einer „stetigen“ Menge von Nebenbedingungen (11) auf eine endliche Anzahl davon begrenzt, die praktisch durch die Möglichkeiten der existierenden Digitalrechnerprogramme bestimmt ist. Die Werte  $t$  können im Intervall  $\langle 0, \pi \rangle$  z. B. gleichmässig verteilt werden.

Da nun mit einem Punkt  $\{\mu_0, \dots, \mu_n\}$ , für den die Bedingung (11) oder (12) erfüllt ist, auch jeder Punkt  $\{a\mu_0, \dots, a\mu_n\}$  für  $a \geq 0$  dieselbe erfüllt, kann man bei einer endlichen Anzahl von Nebenbedingungen (11) von einer konvexen abgeschlossenen „Pyramide“ sprechen, in der die Menge  $M$  enthalten ist.

Da auch hier der zu  $\{\lambda_0, \dots, \lambda_n\}$  nächste Punkt auf der Grenze der „Pyramide“ liegt, so bekommt man als Lösung der vereinfachten Aufgabe „fast sicher“ kein Fejérsches Polynom.

Die Anpassung an ein genaues Fejérsches Polynom, welches (der Stetigkeit wegen) eine Annäherung zur Lösung der ursprünglichen Aufgabe ist, wird durch eine „kleine“ Verschiebung der Wurzeln erreicht.

Bekanntlich [2] muss jede Wurzel eines Fejérschen Polynoms, die auf dem Einheitskreis  $C_1$  liegt, eine gerade Vielfachheit besitzen, und zu jeder Wurzel  $z$  im Inneren von  $C_1$  muss eine Wurzel  $1/\bar{z}$  gehören. Die Wurzel  $z = 0$  ist durch Definition der Fejérschen Polynome ausgeschlossen.

Wir werden bald zeigen, es wird sich bei der Wurzelverschiebung um die Wurzeln auf  $C_1$  handeln.

Die Begründung der beschriebenen anschaulichen Methode beruht auf einigen einfachen Sätzen, die wir jetzt beweisen wollen.

Wir werden das Intervall  $\langle 0, \pi \rangle$  durch sukzessives Halbieren teilen und zu jeder Teilung die zugehörige konvexe „Pyramide“ konstruieren.

**Satz 5.** *Bei dem sukzessiven Halbieren bilden diejenigen Punkte der „Pyramiden“, die dem Punkt  $\{\lambda_0, \dots, \lambda_n\}$  am nächsten liegen, eine beschränkte Folge; die Folge deren Entfernungen vom Punkt  $\{\lambda_0, \dots, \lambda_n\}$  ist nichtabnehmend und beschränkt.*

**Beweis.** Der Koordinatenursprung  $\{0, \dots, 0\}$  liegt in allen „Pyramiden“. So liegen alle dem Punkt  $\{\lambda_0, \dots, \lambda_n\}$  nächsten Punkte der „Pyramiden“ in der „Kugel“ mit dem Mittelpunkt  $\{\lambda_0, \dots, \lambda_n\}$ , welche den Koordinatenursprung enthält. Daraus ist die Beschränktheit sowie der Punktfolge als auch der Folge der Entfernungen von  $\{\lambda_0, \dots, \lambda_n\}$  ersichtlich.

Sollte diese letztgenannte Folge an irgendeiner Position abnehmend sein, so würde das bedeuten, dass es im Inneren einer „Pyramide“ einen Punkt gäbe, dessen Entfernung von  $\{\lambda_0, \dots, \lambda_n\}$  kleiner wäre als die Entfernung des zu  $\{\lambda_0, \dots, \lambda_n\}$  nächsten Punktes der „Pyramide“. Da dieser auf der Grenze der „Pyramide“ liegt, ist das unmöglich.

Somit ist der Satz 5 bewiesen.

**Satz 6.** Die Folge der zu  $\{\lambda_0, \dots, \lambda_n\}$  nächsten Punkte der sukzessiven „Pyramiden“ hat einen einzigen Limespunkt. Dieser ist der zu  $\{\lambda_0, \dots, \lambda_n\}$  nächste Punkt des Durchschnitts der „Pyramiden“.

*Beweis.* Aus dem Satz 5 ist es klar, dass die Folge der zu  $\{\lambda_0, \dots, \lambda_n\}$  nächsten Punkte der sukzessiven „Pyramiden“ mindestens einen Häufungspunkt besitzt und dass die Entfernung aller Häufungspunkte von  $\{\lambda_0, \dots, \lambda_n\}$  dieselbe Limesentfernung ist.

Wir bilden jetzt die zu einem Häufungspunkt konvergierende Unterfolge von Punkten, die die zu  $\{\lambda_0, \dots, \lambda_n\}$  nächsten Punkte der zugehörigen Unterfolge von „Pyramiden“ sind. Da aber alle „Pyramiden“ der Folge ineinander geschachtelt sind, so ist der Durchschnitt jeder Unterfolge mit dem Durchschnitt der ganzen Folge identisch. Der Häufungspunkt liegt also in diesem Durchschnitt.

Weiter ist es klar, dass kein Punkt des Durchschnitts eine kleinere als die Limesentfernung von  $\{\lambda_0, \dots, \lambda_n\}$  haben kann. Gäbe es nun zwei verschiedene Häufungspunkte, so hätte der Durchschnitt, der konvex und abgeschlossen ist, zwei verschiedene zu  $\{\lambda_0, \dots, \lambda_n\}$  nächste Punkte. Das ist unmöglich und der Satz 6 ist bewiesen.

**Satz 7.** Der Durchschnitt  $M$  aller durch (11) definierten Mengen ist mit dem Durchschnitt der bereits definierten „Pyramidenfolge“ identisch.

*Beweis.* Betrachten wir ein  $t \in \langle 0, \pi \rangle$ , das in der sukzessiven Halbierung nicht enthalten ist, und die zu ihm zugehörige Bedingung (11). Da man in der Halbierung eine zu  $t$  konvergente Folge finden kann, so kann man die Bedingung für  $t$  mit der Folge der diesbezüglichen Bedingungen ersetzen. Somit ist der Satz 7 bewiesen.

Mit den Sätzen 6 und 7 haben wir also die Konvergenz der Verfahren des quadratischen Programmierens mit dem steigenden „Auflösungsvermögen“ des Digitalrechners bewiesen.

Da bei jedem Rechner nur eine gewisse Anzahl der Nebenbedingungen bearbeitet werden kann, so bleibt uns noch die Endanpassung an das Fejérsche Polynom durchzuführen, das die Lösung der Approximationsaufgabe ist.

Wir wählen dazu eine Modifikation der Gradientenmethode, die auch selbständig zur Lösung der ganzen Annäherung benutzt werden könnte.

**Satz 8.** Ein Punkt  $\{\mu_0, \dots, \mu_n\}$ , der auf der Grenze von  $M$  liegt, ist genau dadurch charakterisiert, dass das Polynom (1) mit den Koeffizienten  $\mu_0, \dots, \mu_n$  für mindestens ein  $t \in \langle 0, \pi \rangle$  den Wert 0 erreicht.

Beweis. Sollte  $f_n(t)$  im Intervall  $\langle 0, \pi \rangle$  stets positiv sein, so gäbe es in  $\langle 0, \pi \rangle$  ein  $t$ , für welches  $f_n(t)$  den kleinsten Wert  $m > 0$  erreichen würde. Wegen

$$(15) \quad |\cos kt| \leq 1 \quad \text{für } k = 1, \dots, n$$

ergeben die Ungleichheiten

$$(16) \quad \mu_k - \frac{m}{2(n+1)} < \mu_k^* < \mu_k + \frac{m}{2(n+1)} \quad (k = 0, 1, \dots, n)$$

eine Umgebung von  $\{\mu_0, \dots, \mu_n\}$ , die im Inneren von  $M$  liegt. Das ist unmöglich. Somit ist der Satz 8 bewiesen.

Dieser Satz bedeutet also, dass die zu (1) zugehörige reziproke Gleichung

$$(17) \quad \mu_n z^n + \mu_{n-1} z^{n-1} + \dots + \mu_1 z + 2\mu_0 + \mu_1 \frac{1}{z} + \dots + \mu_n \frac{1}{z^n} = 0$$

im Falle eines „Grenzpolynoms“ mindestens eine Doppelwurzel auf dem Einheitskreis  $C_1$  besitzt.

Wir wollen jetzt die Struktur der „Grenzpolynome“ in einfachsten Fällen  $n = 1, 2, 3$  zeigen. Für grössere  $n$  wächst der Rechenaufwand zur Gewinnung der Struktur der „Grenzpolynome“ beträchtlich, die Berechnungen bleiben aber ganz elementar.

#### 6. DER FALL $n = 1$

Die Gleichung (17) ist vom zweiten Grade. Also gibt es genau eine Doppelwurzel auf  $C_1$ , und sie ist entweder  $+1$  oder  $-1$ .

Normieren wir wie Fejér  $\mu_0 = 1$ , so ist  $\mu_1 = \pm 1$ . Die zwei einzigen Fejérschen „Grenzpolynome“ sind also

$$(18) \quad f_1(t) = 1 \pm \cos t.$$

Dazu gehören die Polynome

$$(19) \quad g_1(z) = z \pm 1.$$

Da, wie wir schon bemerkt haben, mit jedem „Grenzpunkt“  $\{\mu_0, \dots, \mu_n\}$  die ganze ihn und den Koordinatenursprung enthaltende Halbgerade zur Grenze von



464  $M$  gehört, so ist hier die Menge  $M$  die durch

$$(20) \quad |\mu_1| = \mu_0$$

begrenzte Viertelebene.

Nimmt man die physikalische Realisierbarkeit in Betracht, so bekommt man zu (19)

$$(21) \quad F(z) = 1 \pm \frac{1}{z}$$

(s. [4], S. 156). Das letztere Filter stellt die Differenz dar und wird in der Radartechnik zur Festzielechounterdrückung benutzt.

DER FALL  $n = 2$

Die Gleichung (17) ist vom vierten Grade. Also gibt es auf der Grenze von  $M$  folgende Wurzelstruktur:

a) es gibt eine Doppelwurzel  $\exp(it)$  auf  $C_1$  und die zugehörige komplex konjugierte Doppelwurzel  $\exp(-it)$ ,

b) es gibt die Doppelwurzel  $+1$  und zwei reziproke reelle Wurzeln  $r$  und  $1/r$ ,

c) es gibt die Doppelwurzel  $-1$  und zwei reziproke reelle Wurzeln  $r$  und  $1/r$ .

In den Punkten  $\pm 1$  schliessen sich die Alternativen nicht einander aus, für das Weitere ist es aber belanglos.

Aus (17) bekommt man dazu (mit der Normierung  $\mu_2 = 1$ )

$$(22) \quad \begin{array}{l} \text{a)} \quad \mu_0 = 2 \cos^2 t + 1, \\ \quad \mu_1 = -4 \cos t, \quad (t \in \langle 0, \pi \rangle) \\ \quad \mu_2 = 1, \\ \\ \text{b)} \quad \mu_0 = R + 1, \quad (R = r + 1/r, \\ \quad \mu_1 = -R - 2, \quad R \geq 2 \text{ oder} \\ \quad \mu_2 = 1, \quad R \leq -2) \\ \\ \text{c)} \quad \mu_0 = -R + 1, \quad (R = r + 1/r, \\ \quad \mu_1 = -R + 2, \quad R \geq 2 \text{ oder} \\ \quad \mu_2 = 1, \quad R \leq -2). \end{array}$$

Fällt in b) und c)  $\mu_0 < 0$  aus, so sind die Vorzeichen von  $\mu_0, \mu_1, \mu_2$  zu wechseln.

Normiert man wieder wie Fejér  $\mu_0 = 1$ , so bekommt man

$$(23) \quad \text{a) } \begin{aligned} \mu_0 &= 1, \\ \mu_1 &= -4 \cos t / (2 \cos^2 t + 1), \quad (t \in \langle 0, \pi \rangle) \\ \mu_2 &= 1 / (2 \cos^2 t + 1), \end{aligned}$$

$$\text{b) } \begin{aligned} \mu_0 &= 1, \\ \mu_1 &= -1 - 1 / (R + 1), \\ \mu_2 &= 1 / (R + 1), \end{aligned}$$

(also  $\mu_2 = -1 - \mu_1$  für  $t = 0$ ),

$$\text{c) } \begin{aligned} \mu_0 &= 1, \\ \mu_1 &= 1 - [1 / (R - 1)], \\ \mu_2 &= -1 / (R - 1), \end{aligned}$$

(also  $\mu_2 = \mu_1 - 1$  für  $t = \pi$ ).

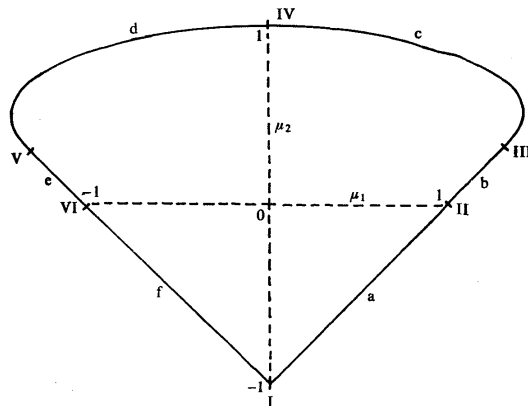


Bild 1.

(23) a) b) c) ergeben zusammen eine Niveaukurve (für  $\mu_0 = 1$ ) der Grenze von  $M$  (s. Bild 1).

Die Grenze selbst ist von Halbgeraden gebildet, die durch den Koordinatenursprung und durch die Punkte der Niveaukurve gehen. Also ist die Grenze aus einem Teil eines Kegels und einem Teil einer Pyramide gebildet. Der gemeinsame Gipfel-punkt liegt im Koordinatenursprung.

466 Im Bild 1 sind die bedeutenden Punkte der Niveaukurve numeriert und die Abschnitte zwischen diesen mit Buchstaben bezeichnet. Im Bild 2 ist die Wurzelnlage der zugehörigen Grenzpolynome skizziert.

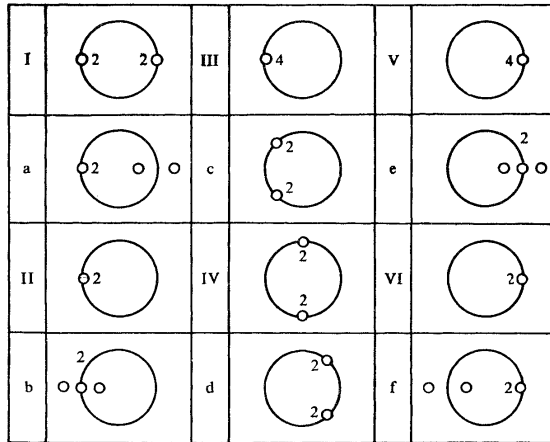


Bild 2.

Unter diesen Grenzpolynomen findet man z. B. das Filter zur Festzielechounterdrückung mit zweifacher Subtraktion sowie ähnliche allgemeinere Filter von Capon (s. [5], [6]).

#### 8. DER FALL $n = 3$

Die Gleichung (17) ist vom sechsten Grade. Also gibt es auf der Grenze von  $M$  folgende Wurzelstruktur:

- a) es gibt eine Doppelwurzel  $\exp(it)$  auf  $C_1$  und die zugehörige komplex konjugierte Doppelwurzel  $\exp(-it)$  und zwei reziproke reelle Wurzeln  $r$  und  $1/r$ ,
- b) es gibt die Doppelwurzel  $+1$  und die vier Wurzeln  $z, \bar{z}, 1/z, 1/\bar{z}$ ,
- c) es gibt die Doppelwurzel  $-1$  und die vier Wurzeln wie in b),
- d) es gibt die Doppelwurzel  $+1$  und die vier reellen Wurzeln  $r_1, r_2, 1/r_1, 1/r_2$ ,
- e) es gibt die Doppelwurzel  $-1$  und die vier Wurzeln wie in d).

Die Alternativen schliessen sich nicht überall aus, für das Weitere ist es aber belanglos.

Aus (17) bekommt man dazu (mit der vorläufigen Normierung  $\mu_3 = 1$ )

$$(24) \quad \begin{aligned} \text{a)} \quad & \mu_0 = -(2 \cos^2 t + 1) R - 4 \cos t, \\ & \mu_1 = 4 \cos^2 t + 4R \cos t + 3, \\ & \mu_2 = -4 \cos t - R, \\ & \mu_3 = 1, \\ \text{b)} \quad & \mu_0 = -R^2 - 2R \cos t - 2 \cos 2t, \\ & \mu_1 = R^2 + 4R \cos t + 2 \cos 2t + 1, \\ & \mu_2 = -2R \cos t - 2, \\ & \mu_3 = 1. \\ \text{c)} \quad & \mu_0 = R^2 - 2R \cos t + 2 \cos 2t, \\ & \mu_1 = R^2 - 4R \cos t + 2 \cos 2t + 1, \\ & \mu_2 = -2R \cos t + 2, \\ & \mu_3 = 1, \\ \text{d)} \quad & \mu_0 = -R_1 R_2 - R_1 - R_2 - 2, \\ & \mu_1 = R_1 R_2 + 2R_1 + 2R_2 + 3, \\ & \mu_2 = -R_1 - R_2 - 2, \\ & \mu_3 = 1, \\ \text{e)} \quad & \mu_0 = R_1 R_2 - R_1 - R_2 + 2, \\ & \mu_1 = R_1 R_2 - 2R_1 - 2R_2 + 3, \\ & \mu_2 = -R_1 - R_2 + 2, \\ & \mu_3 = 1. \end{aligned}$$

Die Bedeutung von  $t$ ,  $R$ , bzw.  $R_1$ ,  $R_2$ , ist klar. Fällt in (24)  $\mu_0 < 0$  aus, so sind alle Vorzeichen von  $\mu_0, \mu_1, \mu_2, \mu_3$  zu wechseln.

## 9. LÖSUNG DER APPROXIMIERUNGSAUFGABE MITTELS DER GRADIENTENMETHODE

**Satz 9.** *Liegt der Punkt  $\{\lambda_0, \dots, \lambda_n\}$  ausserhalb  $M$ , so liegt der Punkt  $\{\mu_0, \dots, \mu_n\}$  auf der Grenze von  $M$ , in dem (13) erreicht wird, zugleich auf einer Geraden, die den Koordinatenursprung enthält, ganz in der Grenze von  $M$  liegt, und mit*

der Verbindungsgeraden der Punkte  $\{0, \dots, 0\}$  und  $\{\lambda_0, \dots, \lambda_n\}$  den kleinsten Winkel bildet (von allen Geraden, die in der Grenze von  $M$  liegen).

Beweis. (13) bedeutet eine „Kugel“ mit dem Mittel  $\{\lambda_0, \dots, \lambda_n\}$ , welche die Grenze von  $M$  im Punkt  $\{\mu_0, \dots, \mu_n\}$  berührt. Auch liegt die ganze Gerade, die die Punkte  $\{0, \dots, 0\}$  und  $\{\mu_0, \dots, \mu_n\}$  verbindet, in der Grenze von  $M$ . Sollte nun eine andere, den Koordinatenursprung enthaltende und in der Grenze von  $M$  liegende Gerade einen kleineren Winkel mit der Verbindung von  $\{0, \dots, 0\}$  und  $\{\lambda_0, \dots, \lambda_n\}$  bilden, so müsste die „Kugel“ diese (und die Grenze von  $M$ ) in einem näheren Punkt berühren. Das ist unmöglich. Dadurch ist der Satz 9 bewiesen.

Setzen wir voraus, dass man die Approximierungsaufgabe für eine „ziemlich feine Pyramide“ durch eine der Standardmethoden der quadratischen Programmierung bereits gelöst hat. Die Lösung

$$(25) \quad \mu_0^* + \mu_1^* \cos t + \dots + \mu_n^* \cos nt$$

ist im Allgemeinen kein Fejérsches Polynom.

Da aber die zu (25) zugehörige Gleichung (17) eine reziproke ist und reelle Koeffizienten besitzt, gehören zu jeder Wurzel  $z$  innerhalb  $C_1$  auch die Wurzeln  $\bar{z}$ ,  $1/z$  und  $1/\bar{z}$ . Der Unterschied von einem Fejérschen Polynom kann also nur darin bestehen, dass die Gleichung (17) mindestens eine Wurzel auf  $C_1$  besitzt, die keine gerade Vielfachheit hat. Man kann aber hoffen, dass es „benachbarte“ Paare von solchen Wurzeln gibt, sodass man sie annähernd durch „Fejérsche“ Wurzeln von doppelter Vielfachheit ersetzen kann.

So hat man zugleich die passende Alternative und die approximativen Werte der Parameter  $t_i$  und  $R_j$  gefunden.

Man wird nun folgendermassen verfahren:

- a) man wählt in der Umgebung des „Parameterpunktes“  $t_i$ ,  $R_j$  ein Netz von Punkten,
- b) man berechnet nach den Formeln der Alternative das Netz von Werten  $\{\mu_0, \dots, \mu_n\}$ ,
- c) man berechnet daraus das Netz von Richtungskosinen

$$(26) \quad x_l = \frac{\mu_l}{\sqrt{\left(\sum_{j=0}^n \mu_j^2\right)}} \quad (l = 0, 1, \dots, n),$$

- d) mit diesen Werten berechnet man

$$(27) \quad v = \lambda_0 x_0 + \dots + \lambda_n x_n,$$

- e) nach einer numerischen Methode für das Maximumsuchen sucht man das Maximum von (27) aus, das dem kleinsten Winkel des Satzes 9 entspricht,

f) man berechnet die „definitiven“ Koeffizienten des gesuchten Fejérschen Polynoms

$$(28) \quad \mu_l = v \cdot \varkappa_l$$

( $l = 0, 1, \dots, n$ , die  $\varkappa_l$  und  $v$  aus dem Maximum von (27)).

Man kann die Methode dieses Absatzes auch selbständig zur Lösung der Fejérschen Annäherung benutzen, man muss jedoch die lokalen Maxima (27) in allen Alternativen berechnen. Das grösste davon stellt die Lösung des Problems dar. Die Existenz der Maxima ist durch die Konvexität der Menge  $M$  gewährleistet.

## 10. BEISPIELE

**Beispiel 1.** Es sei gegeben

$$(29) \quad T_2(t) = 1 + \cos t - \cos 2t.$$

Man überzeugt sich leicht, dass es nicht überall nichtnegativ ist. Durch Anwendung der Methode des Absatzes 9 findet man (annähernd)

$$(30) \quad \mu_0 = 1,33, \quad \mu_1 = 0,67, \quad \mu_2 = -0,67.$$

Das zugehörige Fejérsche trigonometrische Polynom ist

$$(31) \quad f_2(t) = 1,33 + 0,67 \cos t - 0,67 \cos 2t.$$

Im Bild 3 ist durch die vollausgezogene Kurve der Verlauf von (29) gezeigt, die Kreise zeigen die Fejérsche Approximation (31).

Aus der reziproken Gleichung zu (29)

$$(32) \quad z^2 - z - 2 - \frac{1}{z} + \frac{1}{z^2} = 0$$

bekommt man die Wurzeln

$$(33) \quad z_1 = 2,08; \quad z_2 = 0,48, \\ z_{3,4} = -0,78 \pm i \cdot 0,625.$$

Die Fejérschen Wurzeln sind

$$(34) \quad z_1 = 2,62; \quad z_2 = 0,38, \\ z_{3,4} = -1 \quad (\text{Doppelwurzel}).$$

Das zugehörige Polynom (mit den Wurzeln in dem Einheitskreise) ist

$$(35) \quad g_2(z) = 0,82(z + 1)(z - 0,38).$$

Vergleicht man die reellen Wurzeln in (33) mit denen in (34), so sieht man, dass unser Beispiel zugleich ein Gegenbeispiel zur folgenden Vermutung ist: bei der Annäherung könnten die reellen reziproken Wurzeln der anzunähernden Funktion (die ja auch die Wurzeln des Fejérschen trigonometrischen Polynoms vorstellen könnten) unverändert bleiben und die Annäherung würde nur die Lage der Wurzeln verändern, die dem Fejérschen trigonometrischen Polynom nicht angehören können (wie in (33) die zwei weiteren Wurzeln).

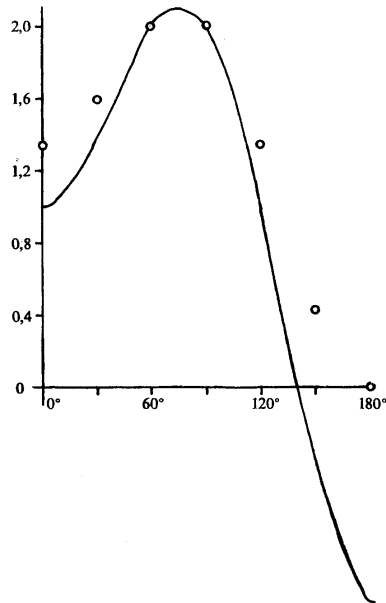


Bild 3.

**Beispiel 2.** Es sei gegeben

$$(36) \quad h(t) = \begin{cases} 1 & \text{für } t \in \langle -l, l \rangle \quad (0 < l < \pi), \\ 0 & \text{im Rest von } \langle -\pi, \pi \rangle. \end{cases}$$

Die zugehörige bekannte Fourierreiheentwicklung ist

$$(37) \quad h(t) \sim \frac{2l}{\pi} \left( \frac{1}{2} + \sum_{n=1}^{\infty} \frac{\sin nl}{nl} \cos nt \right)$$

und stellt mit der Ausnahme der Sprungstellen die Funktion  $h(t)$  dar.

Wählen wir  $t = \pi/4$  und beschränken uns in der Klammer rechts auf drei Glieder. Man bekommt (annähernd)

$$(38) \quad T_2(t) = 0,25 + 0,45 \cos t + 0,32 \cos 2t .$$

Man wird sich leicht Überzeugen, dass (38) nicht überall nichtnegativ ist.

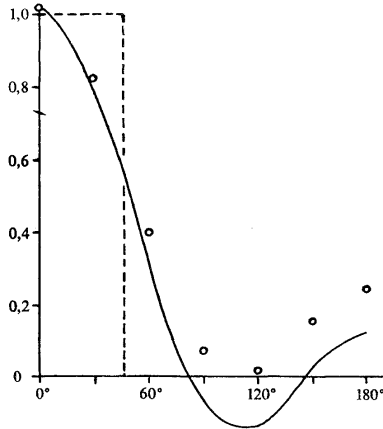


Bild 4.

Durch Anwendung der Methode des Absatzes 9 findet man, dass der der optimalen Annäherung zugehörige Punkt auf dem Kegelmantel liegt mit (annähernd)  $t = 120^\circ$ . Es ist

$$(39) \quad \mu_0 = 0,348 ; \quad \mu_1 = 0,386 ; \quad \mu_2 = 0,282 .$$

Das zugehörige Fejérsche trigonometrische Polynom ist also

$$(40) \quad f_2(t) = 0,348 + 0,386 \cos t + 0,282 \cos 2t .$$

Im Bild 4 ist durch die Strichlinie die Funktion (36) dargestellt. Die vollausgezogene Kurve stellt (38) dar und die Kreise zeigen die Fejérsche Approximation (40).

Aus der reziproken Gleichung zu (38)

$$(41) \quad 0,32z^2 + 0,45z + 0,5 + 0,45 \frac{1}{z} + 0,32 \frac{1}{z^2} = 0$$

bekommt man die Wurzeln

$$(42) \quad \begin{aligned} z_{1,2} &= 0,13 \pm i \cdot 0,99 , \\ z_{3,4} &= -0,81 \pm i \cdot 0,59 . \end{aligned}$$



472 Die Fejérschen Wurzeln sind

$$(43) \quad z_{1,2,3,4} = -0,34 \pm i \cdot 0,94$$

(Doppelwurzeln.)

Das zugehörige Polynom ist

$$(44) \quad g_2(z) = 0,53(z + 0,34 + i \cdot 0,94)(z + 0,34 - i \cdot 0,94).$$

### 11. SCHLUSSBEMERKUNGEN

Da in jeder Alternative der Fejérschen „Grenzpolynome“ immer

$$(45) \quad \mu_l = \varphi_l(\dots, t_l, \dots, R_j, \dots)$$

gilt, so kann man das lokale Minimum (13) in der Alternative auch so suchen, dass man in

$$(13) \quad \Phi(\mu_0, \dots, \mu_n) = \sum_{j=0}^n (\mu_j - \lambda_j)^2$$

$$(46) \quad \mu_l = k\varphi_l$$

einsetzt, wo  $k$  ein unbekannter gemeinsamer Faktor ist, und dann die Lösungen des Systems

$$(47) \quad \begin{aligned} \frac{\partial \Phi}{\partial k} &= 0, \\ &\vdots \\ \frac{\partial \Phi}{\partial t_l} &= 0, \\ &\vdots \\ \frac{\partial \Phi}{\partial R_j} &= 0 \end{aligned}$$

untersucht.

An den Beispielen hat es sich aber gezeigt, dass diese Methode schwieriger anzuwenden ist als diejenige des Absatzes 9.

Jedenfalls ermöglicht die Methode dieses Artikels eine systematische rechnerische Lösung der Aufgabe, die bisher halbempirisch behandelt wurde (s. [4], S. 158).

(Eingegangen am 12. November 1968.)

- [1] Hadley, G.: Nonlinear und dynamic programming. Addison-Wesley, Reading 1964 (rus. Übersetzung: Nelinejnoje i dinamičeskoje programirovanije. Mir, Moskva 1967. Die Zitate im Artikel beziehen sich zu dieser Übersetzung).
- [2] Fejér, L.: Über trigonometrische Polynome. Jour. f. d. reine und angew. Math. (1916), 53—82.
- [3] Prouza, L.: Über die Fejérschen Polynome und einige diskrete Filter und Signale. Kybernetika 5 (1969), 3, 223—232.
- [4] Blackman, R. B.: Linear data smoothing and prediction in theory and practice. Addison-Wesley, Reading, 1965.
- [5] Capon, J.: Optimum weighting functions for the detection of sampled signals in noise. IEEE Trans. *IT-10* (1964), 2, 152—159.
- [6] Robinson, E. A.: Optimum weighting functions for the detection of sampled signals in noise. IEEE Trans. *IT-11* (1965), 3, 452—453.

---

**VÝTAH**

### Aproximace nezáporných funkcí s pomocí Fejérových mnohočlenů jako problém kvadratického programování

LUDVÍK PROUZA

Popisuje se metoda, která — obzvlášť s použitím číslicového počítače — umožňuje aproximovat nezáporné integrovatelné sudé funkce v intervalu  $\langle -\pi, \pi \rangle$  s pomocí Fejérových trigonometrických mnohočlenů ve smyslu nejmenších čtverců.

*Dr Ludvík Prouza, CSc., Ústav pro výzkum radiotechniky, Opočinec, p. Lány na Dálku.*