Werk

Titel: Mathematica

Jahr: 1972

PURL: https://resolver.sub.uni-goettingen.de/purl?312899653_0028|log2

Kontakt/Contact

Digizeitschriften e.V.
SUB Göttingen
Platz der Göttinger Sieben 1
37073 Göttingen

info@digizeitschriften.de

(ACTA F. R. N. UNIV. COMEN. - MATHEMATICA XXVIII 1972)

ACTA FACULTATIS RERUM NATURALIUM UNIVERSITATIS COMENIANAE

MATHEMATICA

PUBL. XXVIII

SLOVENSKÉ PEDAGOGICKÉ NAKLADATEĽSTVO
BRATISLAVA

7

Hlavný redaktor : prof. dr. T. Šalát, CSc

Výkonný redaktor : dr. P. Kostyrko

Redakčná rada

dr. Ing. Jozef Brilla, DrSc, D.Sc
prof. dr. Michal Greguš, DrSc
doc. dr. Milan Hejný, CSc
prof. dr. Anton Huťa, CSc
doc. dr. Tibor Katriňák, CSc
prof. dr. Milan Kolibiar, DrSc
doc. dr. Tibor Neubrunn, CSc
dr. Vladimir Pijak
doc. dr. Milič Sypták, CSc
doc. dr. Valter Šeda, CSc
prof. dr. Marko Švec, DrSc
doc. dr. Štefan Znám, CSc

Austausch von Publikationen erbeten

Priére d´echanger des publicationes

We respectfally solicit the axchange of publications

Se suplica el canje de publicaciónes

(ACTA F. R. N. UNIV. COMEN. - MATHEMATICA XXVIII - 1972)

ACTA FACULTATIS RERUM NATURALIUM UNIVERSITATIS COMENIANAE

MATHEMATICA XXVIII - 1972

MINIMÁLNE POKRYTIE UZAVRETÉHO INTERVALU

OTO ŠTRAUCH, Bratislava

Nech systém I otvorených intervalov pokrýva uzavretý interval $\langle a,b\rangle$. Každý podsystém z I, ktorý pokrýva interval $\langle a,b\rangle$ a ktorý obsahuje najmenší možný počet intervalov z I, nazvime "minimálnym pokrytím uzavretého intervalu $\langle a,b\rangle$ zo systému I". V tejto práci je dokázaný algoritmus, pomocou ktorého vieme nájsť číslo, udávajúce tento minimálny počet intervalov z I potrebných na pokrytie intervalu $\langle a,b\rangle$, a pomocou ktorého vieme z I vybrať každé minimálne pokrytie intervalu $\langle a,b\rangle$. Z tohto algoritmu vyplýva tiež konštruktívny dôkaz Borelovej pokrývacej vety v jednorozmernom prípade.

Úvod

Z Borelovej pokrývacej vety vieme, že ak nejaký systém I otvorených intervalov pokrýva uzavretý interval $\langle a,b\rangle$, potom v I existuje konečný podsystém, ktorý opäť sám pokrýva interval $\langle a,b\rangle$. Preto v I exiistuje tiež taký podsystém, ktorý pokrýva interval $\langle a,b\rangle$ a ktorý obsahuje najmenší možný počet intervalov z I. /Nazvime ho "minimálnym pokrytím uzavretého intervalu $\langle a,b\rangle$ zo systému I"./

Číslo, ktoré udáva tento najmenší možný počet, a každé minimálne pokrytie intervalu $\langle a,b\rangle$ nájdeme pomocou algoritmu, ktorý je uvedený v prvej časti práce.

V druhej časti sú skúmané aditívne vlastnosti minimálneho pokrytia.

Časť prvá

Algoritmus. Nech systém I otvorených intervalov "i" pokrýva uzavretý interval $\langle a,b\rangle$.

Zo systému I vyberme podsystém I_1 všetkých tých intervalov, ktoré obsahujú ľavý koncový bod intervalu $\langle a,b\rangle$, t.j. bod a. Utvorme zjednotenie j_1 všetkých intervalov z I_1. j_1 je otvorený interval.

Zo systému I vyberme podsystém I_2 všetkých tých intervalov, ktoré pretnú interval j_1. Utvorme zjednotenie j_2 všetkých intervalov z I_2.

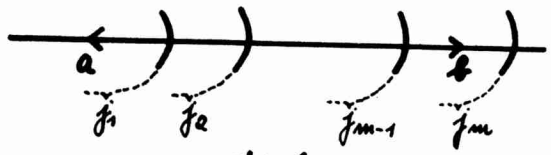
j_2 je otvorený interval.

Zo systému I vyberme podsystém I_3 všetkých tých intervalov, ktoré pretnú interval j_2 a utvorme ich zjednotenie j_3.

j_3 je otvorený interval.

Atď. v n-tom kroku vyberieme zo systému I podsystém I_n všetkých tých intervalov, ktoré pretnú interval j_{n-1} a utvorme ich zjednotenie j_n, čo bude opäť otvorený interval.

<u>V e t a 1.</u> $j_1, j_2, \ldots, j_n, \ldots$ sú otvorené, na seba nadložené intervaly, z ktorých každý obsahuje ľavý koncový bod $\langle a,b\rangle$, t.j. bod a. Po konečnom počte krokov dôjdeme k takému prvému intervalu j_m, ktorý obsahuje i prvý koncový bod intervalu $\langle a,b\rangle$, t.j. bod b. /Teda j_m obsahuje interval $\langle a,b\rangle$ a j_{m-1} neobsahuje interval $\langle a,b\rangle$. /Index "m" intervalu j_m je to hľadané číslo, ktoré udáva minimálny počet intervalov z I, potrebných na pokrytie intervalu $\langle a,b\rangle$.



obr. 1

Ak sme už došli k tomu intervalu j_m, ktorý prvý obsahuje celý interval $\langle a,b\rangle$, pokračujeme ďalej v algoritme takto:

Z podsystému I_m vyberme taký interval i_m, ktorý obsahuje pravý koncový bod intervalu $\langle a,b\rangle$, t.j. bod b.
/To môžeme, keďže $\langle a,b\rangle$ je obsiahnutý v zjednotení j_m intervalov z I_m./

Z podsystému I_{m-1} vyberne taký interval i_{m-1}, ktorý pretne interval i_m.

/To môžeme, pretože i_m je z takého podsystému intervalov, ktoré pretínajú zjednotenie j_{m-1} intervalov z I_{m-1}./

Podobne z I_{m-2} vyberme interval i_{m-2} tak, any preťal i_{m-1} atď., až z I_1 vyberme interval i_1 tak, aby preťal i_2.

V e t a 2. Algoritmom vybraté intervaly $i_1, i_2, \ldots, i_m$ tvoria minimálne pokrytie intervalu $\langle a,b\rangle$ a každé minimálne pokrytie intervalu $\langle a,b\rangle$ možno z I vybrať pomocou tohto algoritmu.

D ô k a z vety 1. Majme teda systém I otvorených intervalov /ktorý pokrýva interval $\langle a,b\rangle$/, z neho algoritmom vybraté podsystémy $I_1, I_2, \ldots, I_n, \ldots$ a zjednotenia intervalov z jednotlivých podsystémov $j_1, j_2, \ldots, j_n, \ldots$ T.j. I_1 obsahuje práve tie intervaly z I, ktoré pretnú interval $\langle a,b\rangle$ v ľavom koncovom bode, atď., až I_n obsahuje práve tie intervaly z I, ktoré pretnú aspoň jeden interval z I_{n-1} alebo, čo je to isté, ktoré pretnú zjednotenie j_{n-1} všetkých intervalov z I_{n-1}. Atď.

Zrejme $I_1 \subset I_2 \subset \ldots \subset I_n \subset \ldots$; $j_1 \subset j_2 \subset \ldots \subset j_n \subset \ldots$

$j_1, j_2, \ldots, j_n, \ldots$ sú otvorené, na seba nadložené intervaly, pretože j_1 je otvorený interval a ak j_{n-1} je otvorený interval, potom tiež j_n je otvorený interval. /Plynie z konštrukcie j_n./

Ak v postupnosti $j_1, j_2, \ldots, j_n, \ldots$ sa dva susedné intervaly rovnajú /$j_k = j_{k+1}$/, potom sa rovnajú i všetky nasledujúce /t.j. $j_k = j_{k+1} = j_{k+2} = \ldots = j_{k+n} = \ldots$/, pretože j_{k+2} je zjednotenie všetkých tých intervalov z I, ktoré pretnú j_{k+1}, ale keďže $j_{k+1} = j_k$, bude j_{k+2} tiež zjednotením všetkých tých intervalov z I, ktoré pretnú j_k. To však je podľa definície j_{k+1}. T.j. $j_{k+1} = j_{k+2}$. Z toho podľa predchádzajúceho plynie $j_{k+2} = j_{k+3}$ atď.

Označme zjednotenie intervalov $j_1, j_2, \ldots, j_n, \ldots$ ako j. j je otvorený interval, ktorý má túto vlastnosť: Ak nejaký interval z I pretne j, potom je už celkom obsiahnutý v j, pretože musí preťať niektorý z intervalov j_n, a teda je obsiahnutý v j_{n+1}.

V dôsledku toho j obsahuje $\langle a,b\rangle$, pretože j obsahuje ľavý koncový interval $\langle a,b\rangle$ a ak by pravý koncový bod j patril do $\langle a,b\rangle$, potom /pretože I pokrýva $\langle a,b\rangle$ /by v I existoval interval, ktorý obsahuje pravý koncový bod intervalu j, t.j. pretína by j a nebol by obsiahnutý v j, čo je spor.

Keďže j je zjednotenie otvorených, na seba nadložených intervalov $j_1, j_2, \ldots, j_n, \ldots$ musí existovať taký interval j_n, ktorý obsahuje interval $\langle a,b\rangle$. Preto v postupnosti $j_1, j_2, \ldots, j_n, \ldots$ musí existovať tiež taký interval j_m, ktorý prvý obsahuje $\langle a,b\rangle$. Potom $\langle a,b\rangle \not\subset j_{m-1}$ a teda $j_{m-1} \neq j_m$. V dôsledku toho $j_1 \neq j_2 \neq \ldots \neq j_m$, pretože, ako sme už videli, ak sa v postupnosti $j_1, j_2, \ldots, j_n, \ldots$ dva susedné intervaly rovnajú, potom sa rovnajú i všetky nasledujúce.

Z $j_1 \neq j_2 \neq \ldots \neq j_m$ plynie, že $I_1 \neq I_2 \neq \ldots \neq I_m$ a preto /keďže $I_1 \neq \emptyset$ a $I_1 \subset I_2 \subset \ldots \subset I_m$/ systém I nemôže mať menej než "m" intervalov /"m" je index toho intervalu j_m, ktorý prvý obsahuje $\langle a,b\rangle$ /.

Nech I´ je ľubovoľný podsystém z I, pokrývajúci $\langle a,b\rangle$. Tak ako sme z I vybrali v algoritme podsystémy $I_1, I_2, \ldots, I_n, \ldots$, vyberme z I´ podsystémy $I_1', I_2', \ldots, I_n', \ldots$

Nech $j_1', j_2', \ldots, j_n' \ldots$ sú zjednotenia intervalov z jednotlivých podsystémov $I_1', I_2', \ldots, I_n', \ldots$

Potom /podľa toho, čo sme dokázali pre ľubovoľný systém I, pokrývajúci $\langle a,b\rangle$ / v postupnosti $j_1', j_2', \ldots, j_n', \ldots$ existuje taký interval j_m', ktorý prvý obsahuje $\langle a,b\rangle$ a podsystém I´ nemôže mať menej než "m´" intervalov.

Keďže však $I' \subset I$, bude $I_1' \subset I_1$, $I_2' \subset I_2$, ... , $I_n' \subset I_n, \ldots$ a teda $j_1' \subset j_1$, $j_2' \subset j_2$, ... , $j_n' \subset j_n$, preto "m´" nemôže byť menšie než "m".

Celkove: Ak podsystém I´ systému I pokrýva interval $\langle a,b\rangle$, potom nemôže mať menej než "m" intervalov.

Na druhej strane, ak dokážeme /v dôkaze vety 2/, že intervaly $i_1, i_2, \ldots, i_m$, vybraté z I /v algoritme/, pokrývajú $\langle a,b\rangle$, bude index "m" intervalu j_m /toho, ktorý prvý z postupnosti $j_1, j_2, \ldots, j_n, \ldots$ obsahuje $\langle a,b\rangle$ /udávať minimálny počet intervalov z I, potrebných na pokrytie $\langle a,b\rangle$. T.j. tým by bol dôkaz vety 1 skončený,

<u>D ô k a z</u> vety 2. Intervaly $i_1, i_2, \ldots, i_m$ sme zo systému I v algoritme vyberali tak, aby sa v susedných dvojiciach pretínali a aby

interval i_1 obsahoval ľavý a interval i_m pravý koncový bod intervalu $\langle a,b \rangle$. Preto zjednotenie intervalov $i_1, i_2, \ldots, i_m$ je opäť otvorený interval, ktorý obsahuje $\langle a,b \rangle$. /Ak sa dva ľubovoľné intervaly pretínajú, potom ich zjednotenie je opäť interval, teda $\cup \{i_k; k = 1,2\}$ je interval, ktorý pretína interval i_3, a teda $\cup \{i_k; k=1,2,3\}$ je interval, ktorý pretína i_4, a teda $\cup \{i_k; k=1,2,3,4\}$ je interval, atď., až $\cup\{i_k; k=1,2,\ldots,m\}$ je interval./ T.j. intervaly $i_1, i_2 \ldots, i_m$ pokrývajú interval $\langle a,b \rangle$ a keďže podľa dôkazu vety 1 ľubovoľné pokrytie intervalu $\langle a,b \rangle$ nemôže mať menej než "m" intervalov, bude $i_1, i_2, \ldots, i_m$ minimálnym pokrytím $\langle a,b \rangle$ v systéme I / a preto intervaly $i_1, i_2, \ldots, i_m$ sú navzájom rôzne./

Ďalej máme dokázať, že ľubovoľné minimálne pokrytie I^o intervalu $\langle a,b \rangle$ v systéme I môžeme z I vybrať použitím tohto algoritmu. T.j., že intervaly "i" z I^o vieme očíslovať od 1 do m tak, aby $i_1 \in I_1$, $i_2 \in I_2, \ldots, i_m \in I_m$, aby sa $i_1, i_2, \ldots, i_m$ v susedných dvojiciach a aby interval i_1 obsahoval ľavý a interval i_m pravý koncový bod intervalu $\langle a,b \rangle$. /"m" je index intervalu j_m, ktorý prvý obsahuje interval $\langle a,b \rangle$, t.j. podľa vety 1 "m" je počet intervalov v I^o./

To môžeme previesť takto: Vyberajme z I^o postupne podľa algoritmu podsystémy $I_1^o, I_2^o, \ldots, I_n^o, \ldots$ a nech $j_1^o, j_2^o, \ldots, j_n^o, \ldots$ sú zjednotenia intervalov z jednotlivých podsystémov. Po konečnom počte krokov dôjdeme k takému intervalu j_m^oo, ktorý ako prvý z postupnosti $j_1^o, j_2^o, \ldots, j_n^o, \ldots$ obsahuje interval $\langle a,b \rangle$. Potom podľa algoritmu vyberajme z podsystémov I_m^oo, $I_{m^o-1}^o, \ldots, I_1^o$ postupne intervaly i_mo, $i_{m^o-1}, \ldots, i_1$ tak, aby $i_m$oobsahoval pravý a i_1 ľavý koncový bod intervalu $\langle a,b \rangle$ a aby sa intervaly i_mo, $i_{m^o-1}, \ldots, i_1$ v susedných dvojiciach pretínali. Podľa prvej časti dôkazu i_mo, $i_{m^o-1}, \ldots, i_1$ je minimálne pokrytie intervalu $\langle a,b \rangle$ v I^o /a preto intervaly i_mo, $i_{m^o-1}, \ldots, i_1$ sú navzájom rôzne/. Keďže však I^o je minimálne pokrytie intervalu $\langle a,b \rangle$ v I, nemôžeme z I^o vypustiť žiaden interval bez toho, že by sme nenarušili pokrytie intervalu $\langle a,b \rangle$. Preto $m=m^o$ a I^o sa skladá práve z intervalov $i_m, i_{m-1}, \ldots, i_1$. Keďže $I^o \subset I$, bude tiež $I_1^o \subset I_1$, $I_2^o \subset I_2, \ldots, I_n^o \subset I_n, \ldots$ a teda $i_1 \in I_1$, $i_2 \in I_2, \ldots, i_m \in I_m$.
Tým je dôkaz vety 2 skončený.

P o z n á m k a 1. Z ľubovoľného systému I otvorených intervalov, ktoré pokrývajú uzavretý interval $\langle a,b \rangle$, vieme vybrať pomocou algoritmu minimálne pokrytie $i_1, i_2, \ldots, i_m$. Pri dôkaze platnosti algo-

ritmu, t.j. pri dôkaze viet 1, 2, sme nepoužili Borelovu vetu. Preto Tento algoritmus predstavuje konštruktívny dôkaz Borelovej pokrývacej vety /v jednorozmernom prípade/.

Poznámka 2. Počet rôznych minimálnych pokrytí intervalu $\langle a,b\rangle$ nie je menší ako počet tých intervalov $i_m \in I_m$, ktoré obsahujú bod b, t.j. môže byť i nespočetne veľký.

Poznámka 3. Za bod, v ktorom sa algoritmus začína, sme mohli zvoliť namiesto ľavého pravý koncový bod intervalu $\langle a,b\rangle$.

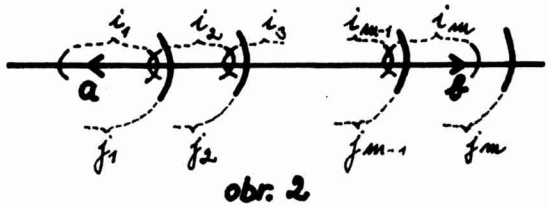
Veta 3. Pre algoritmom vybraté intervaly $i_1, i_2, \ldots, i_m$ platí: $i_2 \in I_2 - I_1$, $i_3 \in I_3 - I_2$, $i_4 \in I_4 - I_3, \ldots, i_m \in I_m - I_{m-1}$ alebo, čo je to isté $a \notin i_2$, $j_1 \cap i_3 = \emptyset$, $j_2 \cap i_4 = \emptyset$, $\ldots$, $j_{m-2} \cap i_m = \emptyset$.

Dôkaz. Majme z I algoritmom vybraté podsystémy $I_1, I_2, \ldots, I_n, \ldots$ Nech $j_1, j_2, \ldots, j_n, \ldots$ sú zjednotenia intervalov z jednotlivých podsystémov. Nech j_m je prvý interval z postupnosti $j_1, j_2, \ldots, j_n, \ldots$, obsahuje interval $\langle a,b\rangle$. Nech $i_1, i_2, \ldots, i_m$ sú algoritmom vybraté intervaly z podsystémov $I_1, I_2, \ldots, I_m$.

Ak by i_k pretínal j_{k-2}, $3 \leqq k \leqq m$ /alebo ak by bod a patril do i_k, $k = 2$/, t.j. ak by i_k patril do I_{k-1}, potom /keďže i_{k+1} pretína i_k /by i_{k+1} patril do I_k a /keďže i_{k+2} pretína i_{k+1}/ i_{k+2} patril do I_{k+1}, atď., až by i_m patril do I_{m-1} a teda by j_{m-1} obsahoval bod b /t.j. aj $\langle a,b\rangle$, čo je spor s definíciou j_m.

Poznámka 1. Algoritmom vybraté intervaly $i_1, i_2, \ldots, i_m$ sa v susedných dvojiciach pretínajú a nepretínajú sa vo vzdialenejších /pretože, ak máme i_k, i_s a $s \leqq k-2$, potom $i_s \subset j_s \subset j_{k-2}$ a $j_{k-2} \cap i_k = \emptyset$ /., t.j. intervaly $i_1, i_2, \ldots, i_m$ sa navzájom oddeľujú.

Poznámka 2. Keďže v algoritme vybratý interval i_k /$3 \leqq k \leqq m$/ nepretína j_{k-2} pretína j_{k-1} a je obsiahnutý v j_k, bude ľavý koncový bod i_k ležať medzi pravými koncovými bodmi j_{k-2} a j_{k-1}, a pravý koncový bod i_k ležať medzi pravými koncovými bodmi j_{k-1} a j_k /podobne pre i_1 a i_2/. Ak ďalej uvážime, že i_k pretne i_{k-1} a i_{k+1}, môžeme algoritmom vybraté intervaly $i_1, i_2, \ldots, i_m$ /a podľa vety 2 aj každé minimálne pokrytie/ zakresliť takto:



obr. 2

Poznámka 3. Každé minimálne pokrytie $\langle a,b\rangle$ môžme z I vybrať pomocou algoritmu práve jedným spôsobom, pretože podľa vety 3 algoritmus prebieha tak, že intervaly $i_m,\ldots, i_2, i_1$ vyberáme postupne z podsystémov $I_m - I_{m-1},\ldots, I_2 - I_1, I_1$, ktoré však /keďže $I_1 \subset I_2 \subset \ldots \subset I_m$/ sú navzájom disjunktné, t.j. každé minimálne pokrytie obsahuje práve po jednom intervale z podsystémov $I_1, I_2 - I_1,\ldots, I_m - I_{m-1}$.

Poznámka 4. Každé minimálne pokrytie môžme očíslovať od 1 do m tak, že $i_1 \in I_1$, $i_2 \in I_2,\ldots, i_m \in I_m$ práve jedným spôsobom, pretože /podľa poznámky 3/ práve jedným spôsobom ho môžme očíslovať tak, aby $i_1 \in I_1$, $i_2 \in I_2 - I_1,\ldots, i_m \in I_m - I_{m-1}$.

/Ak teda máme minimálne pokrytie očíslované tak, že $i_1 \in I_1$, $i_2 \in I_2,\ldots, i_m \in I_m$, potom už platí $i_k \cap i_{k+1} \neq \emptyset$ /k = 1,2,...,m-1/. $a \in i_1, b \in i_m$, t.j. vybrať pomocou algoritmu znamená usporiadať horé uvedeným spôsobom./

Príklad 1. /na prevádzanie algoritmu/ Pokryme interval $\langle a,b\rangle$ tak, že každému bodu x z $\langle a,b\rangle$ priradíme okolie $(x-f(x); x+f(x))$. T.j. pokryme $\langle a,b\rangle$ systémom $I = \{(x-f(x); x+f(x)); x \in \langle a,b\rangle\}$. Nech f(x) je kladná funkcie na $\langle a,b\rangle$ a nech má na $\langle a,b\rangle$ deriváciu. Majme graf funkcie f (x).

Na systéme I prevedieme algoritmus takto:

Zo systému I vyberieme podsystém I_1 tak, že ľavým koncovým bodom intervalu $\langle a,b\rangle$ vedieme priamku p_1 pod uhlom 45°. Všetky okolia /a len tie/, ktoré patria tým bodom x, pre ktoré f(x) leží nad priamku p_1, obsahujú ľavý koncový bod intervalu $\langle a,b\rangle$ a teda tvoria podsystém I_1.

Zjednotenie j_1 okolí z I_1 nájdeme tak, že v grafe f(x) nad priamkou p_1 nájdeme taký bod, v ktorom dotyčnica má smerový uhol -45° a f(x) /v časti nad p_1/ neleží nad touto dotyčnicou. Ak taký bod neexistuje, nájdeme posledný bod, v ktorom f(x) pretína priamku p_1 zhora alebo sa jej zhora dotýka. Ak ani taký bod neexistuje, zoberieme posledný bod grafu funkcie f(x) /t.j. bod s x-ovou súradnicou b/. Týmto bodom vedieme priamku pod uhlom -45°. Bod, v ktorom pretne os x, je pravý koncový bod intervalu j_1.

Zo systému I vyberieme I_2 tak, že pravým koncovým bodom j_1 vedieme priamku p_2 pod uhlom 45°. Všetky okolia /a len tie/, ktoré patria tým bodom x, pre ktoré f(x) leží nad touto priamkou, pretínajú j_1, t.j. tvoria podsystém I_2.

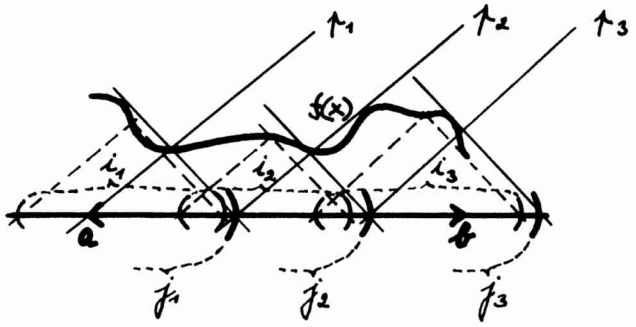
Zjednotenie j_2 okolí z I_2 nájdeme, tak ako v predchádzajúcom prípade, z grafu funkcie f(x) nad priamkou p_2.

Ak sme už, po konečnom počte krokov, došli k prvému intervalu j_m, ktorý obsahuje bod b, pokračujeme v algoritme ďalej takto:

Z I_m vyberieme interval /okolie/ i_m tak, že medzi bodom b a pravým koncovým bodom intervalu j_m nájdeme taký bod, v ktorom priamka vedená pod uhlom -45° pretína graf funkcie f(x) aspoň v jednom bode nad priamkou p_m /stačí hľadať v dostatočnej blízkosti pravého koncového bodu j_m/. V jednom z priesečníkov /nad p_m/ vedieme priamku pod uhlom 45°. Bod, v ktorom táto priamka pretne os x, a pôvodne nájdenný bod, sú krakným bodmi toho hľadaného okolia i_m, ktorý obsahuje bod b a patrí do I_m.

Z I_{m-1} vyberieme okolie i_{m-1} tak, že medzi ľavým koncovým bodom i_m a pravým koncovým bodom j_{m-1} nájdeme bod, v ktorom priamka vedená pod uhlom -45° pretína graf funkcie f(x) aspoň v jednom bode nad priamkou p_{m-1}. V jednom z priesečníkov vedieme priamku pod uhlom 45°. Bod, v ktorom táto priamka pretne os x, a pôvodne nájdenný bod sú krajnými bodmi toho hľadaného okolia i_{m-1}, ktoré pretína i_m a patrí do I_{m-1}. Atď.

V našom prípade /ak f(x) má tvar podľa obrázku/ je minimálny počet intervalov z $I = \{(x-f(x)\,;\ x+f(x))\ ;\ x \in \langle a,b\rangle\}$ potrebný na pokrytie intervalu $\langle a,b\rangle$ rovný 3 a i_1, i_2, i_3 je nejaké minimálne pokrytie.



obr. 3

Č a s ť d r u h á

Nech systém I otvorených intervalov "i" pokrýva uzavretý interval $\langle a,b \rangle$.

Nech x je nejaký bod z intervalu $\langle a,b \rangle$.

Nech m, l_x, p_x sú prirodzené čísla, ktoré /v tomto poradí/ udávajú minimálny počet intervalov z I, potrebných na pokrytie intervalov $\langle a,b \rangle$, $\langle a,x \rangle$ a $\langle x,b \rangle$.

V e t a 4. Pre každý bod x z intervalu $\langle a,b \rangle$ platí: buď $l_x + p_x = m$, alebo $l_x + p_x = m + 1$.

D ô k a z. Vždy je $m \leqq l_x + p_x$.

Na druhej strane, ak $i_1, i_2, \ldots, i_m$ je nejaké, zo systému I algoritmom vybraté minimálne pokrytie intervalu $\langle a,b \rangle$ /t.j. potom $a \in i_1$, $b \in i_m$, $i_k \cap i_{k+1} \neq \emptyset$; $k = 1,2,\ldots,m-1$/ a ak bod x padne do i_k, potom $i_1, i_2, \ldots, i_k$ je nejaké pokrytie intervalu $\langle a,x \rangle$, t.j. $l_x \leqq k$, a $i_k, i_{k+1}, \ldots, i_m$ je nejaké pokrytie intervalu $\langle x,b \rangle$, t.j. $p_x \leqq m-(k-1)$. T.j. $l_x + p_x \leqq m + 1$.

P o z n á m k a. Ak I_1 je nejaké minimálne pokrytie intervalu $\langle a,x \rangle$ a I_2 je nejaké minimálne pokrytie intervalu $\langle x,b \rangle$ /v systéme I/, potom zjednotenie $I_1 \cup I_2$ ešte nemusí byť minimálnym pokrytím intervalu $\langle a,b \rangle$. Ak však pre bod x platí $l_x + p_x = m$, potom vždy je $I_1 \cup I_2$ mini-

málnym pokrytím intervalu $\langle a,b\rangle$ /pretože $I_1 \cup I_2$ pokrýva $\langle a,b\rangle$ a neobsahuje viac než "m" intervalov z I/.

Všetky body x z intervalu $\langle a,b\rangle$, pre ktoré $l_x + p_x = m$, najdeme takto:

Vyberme zo systému I podsystémy $I_1, I_2, \ldots, I_n, \ldots$ tak, že

I_1 je podsystém všetkých tých intervalov z I, ktoré obsahujú bod a,

.
.
.

I_n je podsystém všetkých tých intervalov z I, ktoré pretnú aspoň jeden interval z I_{n-1}.

Atď.

Utvorme zjednotenia $\cup\{i;\ i \in I_1\}$, $\cup\{i;\ i \in I_2\}$, $\cup\{i;\ i \in I_n\}, \ldots$ a označme ich $j_1, j_2, \ldots, j_n, \ldots$

T.j. preveďme na I algoritmus popísaný v prvej časti práce.

Podobne vyberme z I podsystémy $I_1', I_2', \ldots, I_n', \ldots$ tak, že

I_1' obsahuje práve tie intervaly z I, ktoré obsahujú bod b,

.
.
.

I_n' obsahuje práve tie intervaly z I, ktoré pretnú aspoň jeden interval z I_{n-1}'.

Atď.

Utvorme zjednotenia $\cup\{i;\ i \in I_1'\}$, $\cup\{i;\ i \in I_2'\}$, $\cup\{i;\ i \in I_n'\}, \ldots$ a označme ich $j_1', j_2', \ldots, j_n', \ldots$

T.j. preveďme na I algoritmus, ale tak, že namiesto v bode a začneme v bode b.

<u>V e t a 5.</u> množina bodov x z intervalu $\langle a,b\rangle$, pre ktoré platí $l_x + p_x = m$, sa rovná zjednoteniu $\cup\{j_k \cap j'_{m-k};\ k = 1,2,\ldots,m-1\}$.

<u>D ô k a z.</u> Podľa vety 1 vieme, že $l_x = k$ vtedy a len vtedy, ak $x \in j_k$ a $x \notin j_{k-1}$. Podobne $p_x = m-k$ vtedy a len vtedy, ak $x \in j'_{m-k}$ a $x \notin j'_{m-k-1}$.

Teda nutnou podmienkou na to, aby $l_x + p_x = m$ je, aby $x \in j_k \cap j'_{m-k}$ pre nejaké k /$1 \leq k \leq m-1$/.

Na druhej strane uvidíme /veta 7/, že interval j_k nepretína interval j'_{m-k-1} a j_{k-1} nepretína j'_{m-k}. Teda ak x je z prieniku $j_k \cap j'_{m-k}$, potom už x nepatrí ani do j_{k-1} a ani do j'_{m-k-1}, t.j. $l_x = k$ a $p_x = m-k$

Teda to, že $x \in j_k \cap j'_{m-k}$ pre nejaké k /$1 \leq k \leq m-1$/, je aj postačujúcou podmienkou k tomu, aby $l_x + p_x = m$.

V e t a 6. Majme množinu všetkých, zo systému I algoritmom vybratých, minimálnych pokrytí $i_1, i_2, \ldots, i_m$ intervalu $<a,b>$. Potom množina všetkých tých intervalov, ktoré sa v minimálnych pokrytiach $i_1, i_2, \ldots, i_m$ vyskytujú na prvom mieste, je $I_1 \cap I'_m$, na druhom $I_2 \cap I'_{m-1}$ na k-tom $I_k \cap I'_{m-k+1}$, atď., až na m-tom mieste je $I_m \cap I'_1$.

D ô k a z. Majme teda z I vybraté podsystémy $I_1, I_2, \ldots, I_n, \ldots$ /podľa algoritmu, ktorý začína v bode a/ a $I'_1, I'_2, \ldots, I'_n, \ldots$ /podľa algoritmu, ktorý začína v bode b/.

Nech $i_1, i_2, \ldots, i_m$ je nejaké, algoritmom vybraté, minimálne pokrytie intervalu $<a,b>$, /t.j. $i_1 \in I_1$, $i_2 \in I_2, \ldots, i_m \in I_m$, $a \in e_1$, $b \in i_m$, $i_k \cap i_{k+1} \neq \emptyset$; $k = 1,2,\ldots,m-1$/.

Keďže bod b je obsiahnutý v i_m, bude podľa definície $I'_1, i_m \in I'_1$.

Keďže i_{m-1} pretne i_m, bude $i_{m-1} \in I'_2$.

Podobne $i_k \in I'_{m-k+1}, \ldots,$ až $i_1 \in I'_m$.

Teda ak $i_1, i_2, \ldots, i_m$ je algoritmom vybraté minimálne pokrytie intervalu $<a,b>$, potom na m-tom mieste môže byť jedine interval z prieniku $I_m \cap I'_{m-k+1}$.

Na druhej strane nech i je ľubovoľný interval z $I_k \cap I'_{m-k+1}$.

Potom, keďže $i \in I_k$ /a každý interval z I_n pretne aspoň jeden interval v I_{n-1}/, môžme z podsystémov $I_{k-1}, I_{k-2}, \ldots, I_1$ postupne vybrať intervaly $i_{k-1}, i_{k-2}, \ldots, i_1$ tak, aby $i \cap i_{k-1} \neq \emptyset$, $i_{k-1} \cap i_{k-2} \neq \emptyset, \ldots, i_2 \cap i_1 \neq \emptyset$.

Podobne, keďže $i \in I'_{m-k+1}$ /a každý interval z I'_n pretne aspoň jeden interval v I'_{n-1}/, môžeme z podsystémov $I'_{m-k}, I'_{m-k-1}, \ldots, I'_1$ postupne vybrať intervaly $i_{k+1}, i_{k+2}, \ldots, i_m$ tak, aby $i \cap i_{k+1} \neq \emptyset$, $i_{k+1} \cap i_{k+2} \neq \emptyset, \ldots, i_{m-1} \cap i_m \neq \emptyset$.

Tieto intervaly $i_1,\ldots, i_{k-1}, i, i_{k+1},\ldots, i_m$ tvoria minimálne pokrytie intervalu $<a,b>$, pretože je ich práve "m", v susedných dvojiciach sa pretínajú a $a \in i_1$ a $b \in i_m$.

Ďalej interval i nepatrí do I_{k-1}, pretože ak by $i \in I_{k-1}$, potom by sme mohli z podsystémov $I_{k-2}, I_{k-3},\ldots, I_1$ postupne vybrať intervaly $i_{k-2}, i_{k-3},\ldots, i_1$ tak, aby $i \cap i_{k-2} \neq \emptyset$, $i_{k-2} \cap i_{k-3} \neq \emptyset,\ldots, i_2 \cap i_1 \neq \emptyset$. "m - 1" intervalov $i_1,\ldots, i_{k-2}, i, i_{k+1},\ldots, i_m$ by pokrývalo interval $<a,b>$, čo je spor s definíciou "m" /ako čísla udávajúceho minimálny počet intervalov z I, potrebných na pokrytie $<a,b>$/.

Teda $i \in I_k - I_{k-1}$. Podľa poznámky 3 za vetou 3, v každom minimálnom pokrytí existuje práve jeden interval z podsystému $I_k - I_{k-1}$ a ten sa dostane, pri vyberaní tohto minimálneho pokrytia podľa algoritmu, na k-te miesto.

Tým je dôkaz vety 6 skončený.

Poznámka 1. Keďže systém I pokrýva $<a,b>$, existuje aspoň jedno minimálne pokrytie, takže prieniky $I_k \cap I'_{m-k+1}$ /k = 1,2,...,m/, alebo $I_{k+1} \cap I'_{m-k}$ /k = 0,1,..., m-1/ sú neprázdne.

Z $I_{k+1} \cap I'_{m-k} \neq \emptyset$ vyplýva, že podsystém I'_{m-k} obsahuje aspoň jeden interval, ktorý pretína j_k /keďže I_{k+1} je podsystém všetkých tých intervalov z I, ktoré pretnú j_k/ a teda $j'_{m-k} = \cup\{i;\ i \in I'_{m-k}\}$ pretína j_k, t.j. $j_k \cap j'_{m-k} \neq \emptyset$ /k = 1,2,..., m-1/.

T.j. množina bodov x z intervalu $<a,b>$, pre ktoré $l_x + p_x = m$, je pri ľubovoľnom pokrytí I /ak $m > 1$/ neprázdna a keďže je otvorená, bude neprázdny i jej doplnok do $<a,b>$, t.j. bude neprázdna i množina všetkých tých bodov x z $<a,b>$, pre ktoré $l_x + p_x = m + 1$.

Poznámka 2. Ak $i_1, i_2,\ldots, i_m$ je ľubovoľné /z I algoritmom vybraté/ minimálne pokrytie $<a,b>$, potom $i_1 \subset j'_m$, $i_2 \subset j'_{m-1}$, ..., $i_m \subset j'_1$ a $i_k \cap i_{k+1} \subset j_k \cap j'_{m-k}$ /k = 1,2,..., m-1/.

/Pretože $i_k \subset j_k$ a podľa predchádzajúcej vety $i_{k+1} \in I_{k+1} \cap I'_{m-k}$, teda $i_{k+1} \in I'_{m-k}$, t.j. $i_{k+1} \subset j'_{m-k}$/.

V e t a 7. Platí $I_k \cap I'_{m-k} = \emptyset$ /k=1,2,...,m-1/ a tiež $j_k \cap j'_{m-k-1} = \emptyset$ /k = 1,2,...,m-2/.

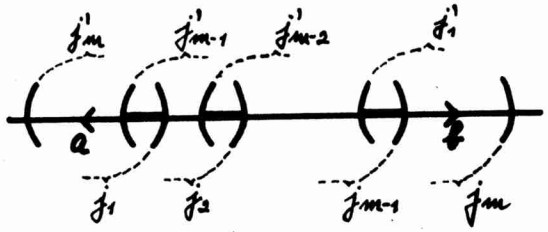
D ô k a z. Interval i_k, ktorý v algoritme vybratom minimálnom pokrytí $i_1, i_2, \ldots, i_m$ stojí na k-tom mieste, je podľa vety 3 z $I_k - I_{k-1}$. Keďže všetky intervaly z $I_k \cap I'_{m-k+1}$ sa vyskytujú na k-tom mieste a $I_{k-1} \subset I_k$, musí byť prienik $I_{k-1} \cap I'_{m-k+1}$ prázdny.

T.j. $I_{k-1} \cap I'_{m-k\ 1} = \emptyset$ /k = 2,3,...,m/, alebo $I_k \cap I'_{m-k} = \emptyset$ /.k=1,2,..,m-1/

Z $I_k \cap I'_{m-k} = \emptyset$ vyplýva, že podsystém I'_{m-k} neobsahuje intervaly, ktoré by pretínali j_{k-1}/t.j. patrili do I_k/ a teda j'_{m-k} nepretína j_{k-1}.

T.j. $j_{k-1} \cap j'_{m-k} = \emptyset$ /k=2,3,..,m-1/, alebo $j_k \cap j'_{m-k-1} = \emptyset$ /k=1,2,...m-2/

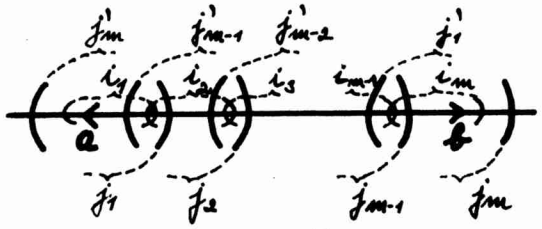
P o z n á m k a. Keďže $j_1 \subset j_2 \subset \ldots \subset j_m$, $j'_1 \subset j'_2 \subset \ldots \subset j'_m$, $a \in j_1$, $a \in j'_m$, $a \notin j'_{m-1}$, $b \in j'_1$, $b \in j_m$, $b \in j_{m-1}$, $j_k \cap j'_{m-k} \neq \emptyset$ /k=1,2,...,m-1/, $j_{k-1} \cap j'_{m-k} = \emptyset$ /k=2,3,...,m-1/, môžeme pravé koncové body intervalov $j_1, j_2, \ldots, j_m$ a ľavé koncové body intervalov $j'_1, j'_2, \ldots, j'_m$ zakresliť takto:



obr. 4

Hrubo vyznačené čiary predstavujú tú množinu bodov x z $\langle a,b \rangle$, pre ktoré $l_x + p_x = m$.

Keďže $i_1 \subset j_1$, $i_2 \subset j_2, \ldots, i_m \subset j_m$, $i_1 \subset j'_m$, $i_2 \subset j'_{m-1}, \ldots, i_m \subset j'_1$, $i_k \cap i_{k+1} \neq \emptyset$ /k = 1,2,...,m-1/, $a \in i_1$, $b \in i_m$, môžeme koncové body ľubovoľného /z I algoritmom vybratého/ minimálneho pokrytia $i_1, i_2, \ldots, i_m$ zakresliť takto:



Obr. 5

Záver

Algoritmus dokázaný v tejto práci môžeme použiť na skúmanie aproximovateľnosti spojitých funkcií schodovitými funkciami na uzavretom intervale $\langle a,b\rangle$, /pod schodovitou funkciou $h(x)$ na uzavretom intervale $\langle a,b\rangle$ rozumieme funkciu, ktorá na $\langle a,b\rangle$ po častiach spojitá a ktorá na $\langle a,b\rangle$ nadobúda iba konečný počet hodnôt. Zúženia $h(x)$ na jednotlivé maximálne intervaly, na ktorých je $h(x)$ konštantná, nazývame schodmi schodovitej funkcie $h(x)$ /pretože ak $f(x)$ je spojitá na intervale $\langle a,b\rangle$, potom pre každé kladné ε môžeme interval $\langle a,b\rangle$ pokryť systémom $I(\varepsilon)$ všetkých takých otvorených okolí $O(x)$ bodov $x \in \langle a,b\rangle$, že pre každé $x' \in O(x)$ platí $|f(x) - f(x')| < \varepsilon$ /pod otvorenými okoliami $O(a)$, $O(b)$ rozumieme otvorené okolia v relatívnej topológii na $\langle a,b\rangle$/a k danej spojitej funkcii $f(x)$ a ku každému kladnému ε vieme nájsť pomocou tohto algoritmu číslo $m(\varepsilon)$, ktoré udáva najmenší možný počet okolí z $I(\varepsilon)$, potrebný na pokrytie intervalu $\langle a,b\rangle$. Pritom platí:

Ak $|f(x) - h(x)| < \varepsilon$ pre každé $x \in \langle a,b\rangle$, potom schodovitá funkcia $h(x)$ nemôže mať menej ako $m(\varepsilon)$ schodov a kù každému kladnému ε axistuje schodovitá funkcia $h(x)$ o $m(\varepsilon)$ schodoch tak, že $|f(x) - h(x)| < \varepsilon$ pre každé $x \in \langle a,b\rangle$.

K spojitej funkcii $f(x)$ existuje na intervale $\langle a,b\rangle$ medzi všetkými schodovitými funkciami s počtom schodov nie väčším ako n funkcia najlepšej aproximácie, t.j. existuje schodovitá funkcia $h^0(x)$ s nie viac ako n schodmi tak, že pre každú schodovitú funkciu $h(x)$ s nie viac ako n schodmi platí $\sup\{|f x - h^0(x)| ; x \in \langle a,b\rangle\} \leq \{\sup |f(x) - h(x)| ; x \in \langle a,b\rangle\}$.

Ak $h^0(x)$ je najlepšou aproximáciou spojitej $f(x)$ spomedzi všetkých nie viac ako n schodových schodovitých funkcií na intervale $\langle a,b\rangle$,

potom rozdiel $|f(x) - h^0(x)|$ nadobúda v niektorých bodoch intervalu $\langle a,b\rangle$ hodnotu $\sup \{|f(x) - h^0(x)| ; x \in \langle a,b\rangle\}$ a $h^0(x)$ má práve $n_0 = \max \{m(\varepsilon) ; m(\varepsilon) \leqq n\}$ schodov.

$h^0(x)$ je najlepšou aproximáciou spojitej $f(x)$ spomedzi všetkých nie viac ako n schodových schodovitých funkcií na intervale $\langle a,b\rangle$ vtedy a len vtedy, keď $\sup \{|f(x) - h^0(x)| ; x \in \langle a,b\rangle\} = \inf \{\varepsilon ; m(\varepsilon) \leqq n\}$.

Ďalej by bolo zaujímavé roztriediť spojité funkcie podľa toho, ako dobre sa dajú aproximovať schodovitými funkciami, t.j. podľa toho, akého rádu je $m(\varepsilon)$ v okolí nuly. Tak napríklad ak $f(x)$ je spojitá a monotónna na intervale $\langle a,b\rangle$ a ak d je dĺžka intervalu $f(\langle a,b\rangle)$ a $[x]$ je celá časť čísla x, potom $m(\varepsilon) = [\varepsilon^{-1}.d] + 1$, teda $m(\varepsilon)$ je v okolí nuly rádove rovné ε^{-1}.

THE MINIMAL COVER OF THE CLOSED INTERVAL

OTO STRAUCH

S u m m a r y

Let the system I of open intervals covers a closed interval $\langle a,b\rangle$.

Let us give a name "the minimal cover of the closed interval $\langle a,b\rangle$ from the system I" to every subsystem of I covering the interval $\langle a,b\rangle$ and containing a minimal number of intervals from I.

In this work an algorithm is proved by means of it we can find a figure assessing the minimal number of intervals from I necessary for the cover of interval $\langle a,b\rangle$ and by means of it we can pick out every minimal cover of the interval $\langle a,b\rangle$ from the system I. It follows from this algorithm a constructive proof of well-konwn Borel covering theorem in one-dimensional case also.

Došlo dňa 31.III.1971 - do redakcie 5.1.1972

Autorova adresa: Laboratórium výpočtovej techniky PFUK, Bratislava
Šmeralova ul. 2/b

(ACTA F. R. N. UNIV. COMEN. - MATHEMATICA XXVIII - 1972)

ACTA FACULTATIS RERUM NATURALIUM UNIVERSITATIS COMENIANAE

MATHEMATICA XXVIII - 1972

BEMERKUNGEN ZU EINEM SYSTEM VON NICHTLINEAREN GLEICHUNGEN

RUDOLF KODNÁR, Bratislava

In dieser Arbeit erwägt man einige Eigenschaften der Gleichungen des Types b/ für flache Schalen und die Lösung des Problems der Bifurkationspunkte. Weiter wird bewiesen, dass das Problem der Stabilität für flache Schalen durch Linearisation der Aufgabe nicht gelöst werden kann.

1.

Differentialgleichungen in der Theorie von schlanken Wänden und flachen Schalen erscheinen in zwei Formen [1] :

a/ als System von drei Gleichungen für Verschiebungen und Durchbiegung.

b/ als System von zwei Gleichungen für die Funktion von Spannung und Durchbiegung.

Die Gleichungen wurden mit Mitteln der numerischen Analyse /grösstenteils ohne Beweis ihrer Verwendbarkeit/ in einer Anzahl von Arbeiten hauptsächlich technischen Charakters untersucht. Insofern uns bekannt ist waren die ersten Arbeiten in welchen die Existenzfragen der Lösung und der Verwendbarkeit einiger direkter Methoden auf die Gleichung des Typs a/ behandelt wurden die Arbeiten [2, 3]. Gleichungen des Typs b/ wurden in [4] erwogen. Später erschienen dann eine immer grössere Menge von Arbeiten in welchen auf verschiedene Art und Weise die Fragen der Existenz und auch der Regularität der Lösungen beider Typen a/ und b/ sowie die Fragen der Verwendbarkeit der Näherungsmethoden und die Art ihrer Konvergenz geprüft wurden. Von diesen erwähnen wir nur [5,6,7,8,9,10]. Die Probleme der Lösbarkeit von Systemen algebraischer Gleichungen welche wir durch Verwendung der direkten Methode auf das System des Typs b/ erhalten wurden z.B. in [2, 11] untersucht.

Das Problem der Stabilität schlanker Wände und flacher Schalen wurde für Gleichungen des Typs a/ zum ersten mal durch begründete Linearisation in [12] gelöst. Für Typ b/ erschien das Existenzproblem mehrerer Zweige der Lösung, das Verschwinden der Stabilen Form der Lösung und der Übergang der Form der Lösung auf eine andere Form in [13] bei der Lösung von Systemen algebraischer Gleichungen, zu welchen die Verwendung der direkten Methode führte.

Einige theoretische Erwägungen dieser Probleme finden wir in [14] Das Problem der Linearisation für Gleichungen schlanker Wänder des Typs b/ ist in [6, 9] für sehr allgemeine Randbedingungen gelöst worden.

2. Grundlegende Begriffe

Wir werden uns mit dem System für w, F partieller Differentialgleichungen der Theorie flacher Schalen folgender Form beschäftigen

$$\frac{D}{h}\Delta\Delta w = \lambda\,[F_0,\, w] + [k,\, F] + [F,\, w]\,, \tag{1}$$

$$-\frac{1}{E}\Delta\Delta F = \frac{1}{2}\,[w,w] + [k,w]\,. \tag{2}$$

In den Gleichungen sind D, h, E positive Konstante, k (x,y) ist eine genügend reguläre Funktion, $F_0(x,y)$ ist die Funktion der Spannung in der Mittelfläche der Schale. Setzen wir voraus, dass $F_0 \not\equiv 0$ und genügend regulär ist. Weiter bezeichnen wir

$$[\alpha,\beta] = \frac{\partial^2\alpha}{\partial x^2}\frac{\partial^2\beta}{\partial y^2} + \frac{\partial^2\alpha}{\partial y^2}\frac{\partial^2\beta}{\partial x^2} - 2\,\frac{\partial^2\alpha}{\partial x\partial y}\frac{\partial^2\beta}{\partial x\partial y}\,.$$

Wir werden nur mit reellen Funktionen w,F arbeiten und λ sei ein reeller Parameter.

Die Gleichungen (1), (2) werden wir im begrenzten Gebiet Ω des zweidimensionalen Euklidischen Raumes ($\Omega \subset E_2$) untersuchen. Die Grenze von Ω werden wir mit $\partial\Omega$ bezeichnen. Die Randbedingungen für (1), (2) können wir durch Kombination dieser

$$\left.\begin{aligned} &a_1/\quad w|_{\partial\Omega} = \frac{\partial w}{\partial n}|_{\partial\Omega} = 0\ ,\quad a_2/\quad F|_{\partial\Omega} = \frac{\partial F}{\partial n}|_{\partial\Omega} = 0, \\ &b_1/\quad w|_{\partial\Omega} = \frac{\partial^2 w}{\partial x^2}|_{\partial\Omega} = \frac{\partial^2 w}{\partial y^2}|_{\partial\Omega} = 0,\quad b_2/\quad F|_{\partial\Omega} = \frac{\partial^2 F}{\partial x^2}|_{\partial\Omega} = \frac{\partial^2 F}{\partial y^2}|_{\partial\Omega} = 0, \end{aligned}\right\} 3$$

bekommen, wo n die Aussennormale zu $\partial\Omega$ ist.

Die schwache Lösung $< w,F >$ des Grundsystems (1),(2) werden wir am Raum $W = H \times H$ definieren. Den Raum H werden wir mit Rücksicht auf die Randbedingungen (3) folgend definieren:

Wenn irgendeine von den Funktionen w,F die Randbedingungen $a_1/$, $a_2/$ /Dirichletsche Aufgabe/ erfüllen soll, dann werden wir voraussetzen, dass die Grenze eine Lipschitzische ist und auf die entsprechende Stelle im kartesischen Produkt für W setzen wir

$H = \overset{o}{W}{}_2^{(2)}(\Omega)$ [16].

Wenn irgendeine von den Funktionen w,F die Bedingungen $b_1/$, $b_2/$ erfüllen soll, werden wir voraussetzen, dass Ω ein rechteckiges Parallelepiped mit den Seiten a,b in $E_2(\Omega \equiv <0,a> x <0,b>)$ ist. Auf bekannte Art bezeichnen wir $C^{(k)}(\bar{\Omega})$ die Menge der Funktionen, welche auf $\bar{\Omega}$ alle Ableitungen bis zur k-ten Ordnung inclusive stetig haben. Wir bezeichnen $D_A \equiv \{w \in C^{(4)}(\bar{\Omega}), b\}$ /die Menge der Funktionen welche auf $\bar{\Omega}$ vier stetige Ableitungen haben und die $b_1/$ Randbedingungen erfüllen/.

Ersichtlich gilt $D_A \subset L_2(\Omega)$, $\bar{D}_A = L_2(\Omega)$.

B e m e r k u n g 1. Es ist zu sehen, dass eine ähnliche D_A auch für die Randbedingungen $a_1/$ existiert.
Es ist / durch Verwendung der Friedrichschen und Poincaréschen Ungleichung [10]/, dass für $w \in D_A$

$$(\Delta\Delta\, w,w)_{L_2(\Omega)} \geq \gamma^2 \int_{\bar{\Omega}} w^2 d\Omega = \gamma^2 \| w \|^2_{L_2(\Omega)} \qquad (4)$$

ist, wo

$$\gamma^2 = \frac{1}{a^2 \cdot \max(a^2, b^2)}$$

Durch die Vorschrift

$$[w,v]_\Delta = (\Delta\Delta w, v)_{L_2(\Omega)} = \int_\Omega \Delta w \, \Delta v \, d\Omega , \qquad w,v \in D_A$$

führen wir auf D_A ein neues skalares Produkt ein.

Wir definieren die Norm

$$\|w\|_0^2 = [w,w]_\Delta , \qquad w \in D_A . \tag{5}$$

Aus (4) ist dann ersichtlich, dass

$$\|w\|_0 \geq \gamma \|w\|_{L_2(\Omega)} . \tag{6}$$

Wenn wir D_A in der Norm /5/ vervollständigen, erhalten wir einen neuen Hilbertschen Raum, welchen wir mit H_0 bezeichnen. Bekannt ist der Raum $W_2^{(2)}(\Omega)$ von Sobolev [17]. $W_2^{(2)}(\Omega)$ ist in der Norm

$$\|w\|^2_{W_2^{(2)}(\Omega)} = \int_\Omega w^2 d\Omega + \int_\Omega \left[\left(\frac{\partial^2 w}{\partial x^2}\right)^2 + 2 \left(\frac{\partial^2 w}{\partial x \partial y}\right)^2 + \left(\frac{\partial^2 w}{\partial y^2}\right)^2 \right] d\Omega \tag{7}$$

vollständig. Ersichtlich ist $D_A \subset W_2^{(2)}(\Omega)$. Aus /6/ und /7/ ist aber auch ersichtlich, dass für $w \in D_A$

$$\|w\|_0 \leq \|w\|_{W_2^{(2)}(\Omega)} \leq \mathcal{H} \|w\|_0 , \qquad \mathcal{H} = \sqrt{a^2 \cdot \max(a^2, b^2) + 1} \tag{8}$$

ist. Die Normen /5/, /7/ sind also auf D_A equivalent und H_0 ist also eine Vervollständigung von D_A in der Norm des Raumes $W_2^{(2)}(\Omega)$. Untersuchen wir jetzt die Eigenschaften des Raumes H_0. Es sei $w \in H_0$. Dann ist auch $w \in L_2(\Omega)$. Ausserdem existiert die Folge $\{w_n\}_1^\infty \subset D_A$ derat, dass $\|w - w_n\|_0 \xrightarrow[n\to\infty]{} 0$ ist. Es gilt also auch

$$\|w_n - w_m\|_0 \xrightarrow[n,m\to\infty]{} 0 \quad \|w_n - w_m\|_{W_2^2} \xrightarrow[n,m\to\infty]{} 0 .$$

Durch Benutzung der Friedrichschen Ungleichung erhalten wir gleich, dass die erste und zweite Ableitung der Funktion w_n in $L_2(\Omega)$ konvergieren. Es ist also ersichtlich /aus den Definitionen der verallgemeinerten Ableitung [17] /dass w alle verallgemeinerten Ableitungen erster und zweiter Ordung besitzt, welche in $L_2(\Omega)$ gehören. Also ist $w \in W_2^{(2)}(\Omega)$. Dies besagt aber nichts anderes, als dass $H_0 \subset W_2^{(2)}(\Omega)$. Als vollständiger Teil des Raumes $W_2^{(2)}(\Omega)$ ist H_0 in $W_2^{(2)}(\Omega)$ abgeschlossen und ist ein Unterraum des Raumes $W_2^{(2)}(\Omega)$. Die Sätze über Einbettung [17] sind also für H_0 verwendbar. Aus der Dichte von D_A in H_0 folgt, dass die Relationen /5/, /8/ auch für $w \in H_0$ wahrheitsgemäss sind. Dabei sind im Operator Δ jetzt schon die Ableitungen verallgemeinert/dass werden wir im Weiternen allerdings nicht speziell bezeichnen/. In W werden wir dann an ensprechender Stelle im kartesischen Produkt $H=H_0$ setzen.

In den Fällen beider Typen von Randwertproblemen aus /3/ werden wir den Raum auf welchen wir eine schwache Lösoung suchen formal ebenso bezeichen. Es ist also $W = H \times H$, wobei $H = \mathring{W}_2^{(2)}(\Omega)$, oder $H = H_0$. Alle Berichtigungen und Ubergänge im Text sind für beide Fälle H verwendbar. Bei Abschätzungen werden wir die Konstanten für den Fall $H = H_0$ berechnen. Die Konstanten für $H = W_2^{(2)}(\Omega)$ folgen aus den allgemeinen Sätzen über Einbettung.

3. Die Eigenschaften der nichthomogenen biharmonischen Gleichung bei den Randbedingungen /3/

Wählen wir die Differentialgleichung $\Delta\Delta F = f$. /9/. Für ein gewisses $f \in C(\Omega)$ sei $F(f) \in D_A$ die Lösung der Gleichung /9/. Für jedes $v \in D_A$ erhalten wir dann aus /9/

$$(\Delta\Delta F, v)_{L_2(\Omega)} = (f, v)_{L_2(\Omega)} ,$$

oder auch

$$[F(f), v]_\Delta = \int_\Omega fv \, d\Omega .$$

Aus der Dichte D_A in H erhalten wir gleich, dass die letzte Gleichheit auch für jedes $v \in H$.

Es sei $F \in H$, dann ist $F \in W_2^{(2)}(\Omega)$ und aus den Sätzen über Einbettung erhalten wir, dass $F \in C(\bar{\Omega})$ und es gilt

$$\max_{\bar{\Omega}} |F| \leq \alpha \|F\|_{W_2^{(2)}(\Omega)} \leq \beta \|F\|_0, \quad \beta = \alpha \mathcal{H}. \qquad /10/$$

Es sei $f \in L_1(\Omega)$, für jedes $v \in H$ gilt dann

$$\left| \int_\Omega f\, v\, d\Omega \right| \leq \left(\max_{\bar{\Omega}} |v| \right) \int_\Omega |f|\, d\Omega \leq \alpha \|v\|_{W_2^{(2)}(\Omega)} \int_\Omega |f|\, d\Omega \qquad /11/$$

<u>D e f i n i t i o n 1</u>. Das Element $F \in H$ werden wir bei gegebenem $f \in L_1(\Omega)$ eine schwache Lösung der Gleichung /9/ nennen, wenn $\forall v \in H$ gilt

$$[F(f), v]_\Delta = \int_\Omega f\, v\, d\Omega . \qquad /12/$$

Aus der Bilinearität der rechten Seite /12/ und aus /11/ folgt, dass $\int_\Omega f\, v\, d\Omega$ für das gegebene $f \in L_1(\Omega)$ ein linear begrenztes Funktional auf H ist. Gemäss dem Satz vo Riesz [18] entspricht dann jedem $f \in L_1(\Omega)$ ein einziges Element $\Psi(f) \in H$ derart, dass

$$\int_\Omega f\, v\, d\Omega = [\Psi(f), v]_\Delta , \quad \forall v \in H \qquad /13/$$

gilt. Die Gleichung /9/ hat also für jedes $f \in L_1(\Omega)$ gerade eine schwache Lösung $\Psi(f) = F(f)$. Aus /11/ und /13/ erhalten wir

$$\|F(f)\|_{W_2^{(2)}(\Omega)} \leq \alpha \int_\Omega |f|\, d\Omega . \qquad /14/$$

Für $w \in H$ und eine genügend glatte Funktion $k(x,y)$ erhalten wir aus /7/, /8/ die Ungleichung

$$\frac{1}{2} \int_\Omega |[w,w] + 2\dot{} [k,w]|\, d\Omega \leq 2 \mathcal{H}^2 \left(\|w\|_0^2 + A \|w\|_0 \right) , \qquad /15/$$

wo A eine entsprechende Konstante ist. Also entspricht jedem $w \in H$ gerade eine schwache Lösung $F \in H$ der Gleichung

$$\frac{1}{E} \Delta\Delta F = -\frac{1}{2}[w,w] - [k,w]$$

und es gilt die Abschätzung

$$\|F(w)\|_0 \leq \varrho \left(\|w\|_0 + A\right) \|w\|_0, \quad w \in H, \ \varrho = 2E\varkappa^2\alpha . \qquad /16/$$

4. Die schwache Lösung der Aufgabe /1/, /2/, /3/

Die Norm für das Paar $\langle w,F \rangle \in W$ definieren wir auf normale Weise

$$\|\langle w,F \rangle\|_W^2 = \|w\|_H^2 + \|F\|_H^2 .$$

Definition 2. Wir werden sagen, dass das Paar $\langle w,F \rangle \in W$ eine schwache Lösung des systems /1/, /2/, bei den Randbedingungen /3/ ist, wenn für jedes Element $\langle \tilde{w},\tilde{F} \rangle \in W$

$$\frac{D}{h}\int_\Omega \Delta w \Delta \tilde{w}\, d\Omega - \frac{1}{E}\int_\Omega \Delta F \Delta \tilde{F}\, d\Omega - \lambda\int_\Omega [F_0,w]\,\tilde{w}\, d\Omega -$$
$$- \int_\Omega [k,F]\,\tilde{w}\, d\Omega - \int_\Omega [F,w]\,\tilde{w}\, d\Omega - \frac{1}{2}\int_\Omega [w,w]\,\tilde{F}\, d\Omega -$$
$$- \int_\Omega [k,w]\,\tilde{F}\, d\Omega = 0$$

gilt. Nach Berichtigung erhalten wir, dass $\langle w,F \rangle \in W$ eine schwache Lösung von /1/, /2/ bei den Randbedingungen /3/ ist, wenn für alle $\langle \tilde{w},\tilde{F} \rangle \in W$ folgendes gilt:

$$A(w,F)\langle\tilde{w},\tilde{F}\rangle \equiv \frac{D}{h}\int_{\Omega}\Delta w\Delta\tilde{w}\,d\Omega - \frac{1}{E}\int_{\Omega}\Delta F\Delta\tilde{F}\,d\Omega -$$

$$- \lambda\int_{\Omega}\Big[\Big(\frac{\partial^2 F_0}{\partial x\partial y}\frac{\partial w}{\partial y} - \frac{\partial^2 F_0}{\partial y^2}\frac{\partial w}{\partial x}\Big)\frac{\partial\tilde{w}}{\partial x} + \Big(\frac{\partial^2 F_0}{\partial x\partial y}\frac{\partial w}{\partial x} -$$

$$- \frac{\partial^2 F_0}{\partial x^2}\frac{\partial w}{\partial y}\Big)\frac{\partial\tilde{w}}{\partial y}\Big]d\Omega - \int_{\Omega}\Big[\Big(\frac{\partial^2 k}{\partial x\partial y}\frac{\partial F}{\partial y} - \frac{\partial^2 k}{\partial y^2}\frac{\partial F}{\partial x}\Big)\frac{\partial\tilde{w}}{\partial x} +$$

$$+ \Big(\frac{\partial^2 k}{\partial x\partial y}\frac{\partial F}{\partial x} - \frac{\partial^2 k}{\partial x^2}\frac{\partial F}{\partial y}\Big)\frac{\partial\tilde{w}}{\partial y}\Big]d\Omega - \int_{\Omega}\Big[\Big(\frac{\partial^2 F}{\partial x\partial y}\frac{\partial w}{\partial y} -$$

$$- \frac{\partial^2 F}{\partial y^2}\frac{\partial w}{\partial x}\Big)\frac{\partial\tilde{w}}{\partial\tilde{x}} + \Big(\frac{\partial^2 F}{\partial x\partial y}\frac{\partial w}{\partial x} - \frac{\partial^2 F}{\partial x^2}\frac{\partial w}{\partial y}\Big)\frac{\partial\tilde{w}}{\partial y}\Big]d\Omega +$$

$$+ \frac{1}{2}\int_{\Omega}\Big(\frac{\partial w}{\partial x}\frac{\partial^2 w}{\partial y^2}\frac{\partial\tilde{F}}{\partial x} - \frac{\partial w}{\partial x}\frac{\partial^2 w}{\partial x\partial y}\frac{\partial\tilde{F}}{\partial y}\Big)d\Omega -$$

$$- \int_{\Omega}\Big[\Big(\frac{\partial^2 k}{\partial x\partial y}\frac{\partial w}{\partial y} - \frac{\partial^2 k}{\partial y^2}\frac{\partial w}{\partial x}\Big)\frac{\partial\tilde{F}}{\partial x} + \Big(\frac{\partial^2 k}{\partial x\partial y}\frac{\partial w}{\partial x} -$$

$$- \frac{\partial^2 k}{\partial x^2}\frac{\partial w}{\partial y}\Big)\frac{\partial\tilde{F}}{\partial y}\Big]d\Omega = 0\,. \qquad (17)$$

5. Die Potentialität des Operators A(w,F)

Es sei B ein reeller Banachscher Raum. G sei ein auf B definierter Operator. t sei eine reelle Zahl.

Definition 3. Wenn im Punkte $\eta \in B$

$$\lim_{t\to 0}\frac{G(\eta + t\xi) - G(\eta)}{t} = V\,G(\eta,\xi), \quad \xi \in B$$

existiert, dann nennen wir den Operator $V\,G(\eta,\xi)$ einen Gateauxschen /oder einen schwachen/ Differentialoperator G im Punkte η. Wenn $V\,G(\eta,\xi)$ in Bezug auf ξ linear ist, bezeichnen wir diesen $D\,G(\eta,\xi)$ und nennen ihn einen linearen Gateaux-schen Differential operator G im Punkte η.

Definition 4. Es sei ϕ ein Funktional auf B, welches auf irgendeiner Menge $M \subset B$ ein lineares Gateaux-sches Differential $D\phi(\eta, \xi)$ hat.

Den durch die Beziehung

$$D\phi(\eta, \xi) = (G(\eta), \xi)$$

definierten Operator $G(\eta)$ nennen wir Gradienten des Funktionals $\phi(\eta)$ und wir schreiben

$$G(\eta) = \text{grad}\, \phi(\eta) .$$

Definition 5. G sei ein Operator von B zu B^* /adjungiert zu B/. Wir sagen, dass G auf der Menge $M \subset B$ potential ist, wenn ein solches Funktional ϕ existiert, dass für alle $\eta \in M$ grad $\phi(\eta) = G(\eta)$ ist.

Hilfssatz 1. ([19], §5). Es sei $G(\eta)$ ein Operator von B zu B^*. Es gelte:

1/ $G(\eta)$ hat ein Gateaux-sches Differential $DG(\eta, \xi)$ in jedem Punkt η der Kugel $D_r = \{\eta : \| \eta - \eta_0 \|_B < r, r > 0\}$.

2/ In jedem Punkt $\eta \in D_r$ ist das Funktional $DG(\eta, \xi), \xi$ mit Rücksicht auf η stetig.

Die notwendige und hinreichende Bedingung dazu, dass $G(\eta)$ auf D_r potential sei, ist

$$DG(\eta, \xi)\xi = DG(\eta, \xi)\xi .$$

Es existiert ein einziges Funktional $\phi(\eta)$, welches in η_0 den gegebenen Wert ϕ_0 erhält, dessen Gradient der Operator $G(\eta)$ ist. Dieses Funktional hat die Form

$$\phi(\eta) = \phi_0 + \int_0^1 (G(\eta_0 + t(\eta - \eta_0)), \eta - \eta_0) dt. \qquad /18/$$

Es gilt folgender

Hilfssatz 2. Es sei $\varphi(x,y)$, $\psi(x,y) \in H$ und es sei $\nu(x,y) \in W_2^{(2)}(\Omega)$. Dann gilt

$$([\nu,\varphi],\psi)_{L_2(\Omega)} = ([\nu,\psi],\varphi)_{L_2(\Omega)} .$$

Beweis

$$\int_\Omega \left(\frac{\partial^2 \nu}{\partial x^2} \frac{\partial^2 \varphi}{\partial y^2} + \frac{\partial^2 \nu}{\partial y^2} \frac{\partial^2 \varphi}{\partial x^2} - 2 \frac{\partial^2 \nu}{\partial x \partial y} \frac{\partial^2 \varphi}{\partial x \partial y} \right) \psi \, d\Omega = \int_\Omega \frac{\partial}{\partial y} \left(\frac{\partial^2 \nu}{\partial x^2} \frac{\partial \varphi}{\partial y} - \right.$$

$$\left. - \frac{\partial^2 \nu}{\partial x \partial y} \frac{\partial \varphi}{\partial x} \right) \psi \, d\Omega + \int_\Omega \frac{\partial}{\partial x} \left(\frac{\partial^2 \nu}{\partial y^2} \frac{\partial \varphi}{\partial x} - \frac{\partial^2 \nu}{\partial x \partial y} \frac{\partial \varphi}{\partial y} \right) \psi \, d\Omega = \int_\Omega \frac{\partial^2}{\partial x^2} \left(\frac{\partial^2 \nu}{\partial y^2} \varphi \right) \psi \, d\Omega$$

$$- 2 \int_\Omega \frac{\partial^2}{\partial x \partial y} \left(\frac{\partial^2 \nu}{\partial x \partial y} \varphi \right) \psi \, d\Omega + \int_\Omega \frac{\partial^2}{\partial y^2} \left(\frac{\partial^2 \nu}{\partial x^2} \varphi \right) \psi \, d\Omega = \int_\Omega \frac{\partial^2 \nu}{\partial y^2} \frac{\partial^2 \psi}{\partial x^2} \varphi \, d\Omega -$$

$$- 2 \int_\Omega \frac{\partial^2 \nu}{x \; y} \frac{\partial^2 \psi}{\partial x \partial y} \varphi \, d\Omega + \int_\Omega \frac{\partial^2 \nu}{\partial x^2} \frac{\partial^2 \psi}{\partial y^2} \varphi \, d\Omega .$$

Satz 1. Der Operator A (w,F) in /17/ ist potential auf W.

Beweis. Zeigen wir, dass beide Voraussetzungen des Hilfssatzes 1 erfüllt sind. Bezeichnen wir $\eta = \langle w,F \rangle$, $\xi = \langle \tilde{w},\tilde{F} \rangle$, $\zeta = \langle \bar{w},\bar{F} \rangle \in H$. Dann hat /17/ die Form

$$A(\eta)\xi = 0. \qquad /19/$$

konstruieren wir ein Gateaux-sches Differential /19/ (wenn dies möglich ist), berechnen wir also den Quotient

$$\frac{A(\eta + t\zeta)\xi - A(\eta)\xi}{t} = \frac{1}{t} \left\{ \frac{D}{h} \left(\Delta(w+t\bar{w}) \; , \Delta\tilde{w} \right)_{L_2(\Omega)} - \right.$$

$$- \frac{1}{E} \left(\Delta(F+t\bar{F}) , \Delta\tilde{F} \right)_{L_2(\Omega)} - \lambda \left([F_0, w+t\bar{w}] \; , \tilde{w} \right)_{L_2(\Omega)} -$$

$$- \left([k, F+t\bar{F}], \tilde{w} \right)_{L_2(\Omega)} - \left([F+t\bar{F}, w+t\bar{w}] \; , \tilde{w} \right)_{L_2(\Omega)} -$$

$$\left. - \frac{1}{2} \left([w+t\bar{w}, w+t\bar{w}], \tilde{F} \right)_{L_2(\Omega)} - \left([k, w+t\bar{w}] \; , \tilde{F} \right)_{L_2(\Omega)} - A(\eta)\xi \right\} .$$

Durch Anwendung des Hilfssatzes 2 können wir wie folgt schreiben

$$\frac{A(\eta+t\xi)\xi - A(\eta)\xi}{t} = \frac{1}{t}\left\{\frac{D}{h}\left[(\Delta w,\Delta\tilde{w})_{L_2(\Omega)} + t(\Delta\bar{w},\Delta\tilde{w})_{L_2(\Omega)}\right] -\right.$$

$$-\frac{1}{E}\left[(\Delta F,\Delta\tilde{F})_{L_2(\Omega)} + t(\Delta\bar{F},\Delta\tilde{F})_{L_2(\Omega)}\right] - \lambda\left([F_0,\tilde{w}],\ F+t\bar{F}\right)_{L_2(\Omega)} -$$

$$-\left([k,\tilde{w}],\ F+t\bar{F}\right)_{L_2(\Omega)} - \left([F+t\bar{F},\tilde{w}],\ w+t\tilde{w}\right)_{L_2(\Omega)} - \frac{1}{2}\left([w+t\bar{w},\tilde{F}],\ w+t\bar{w}\right)_{L_2(\Omega)} -$$

$$\left.-\left([k,\tilde{F}],\ w+t\bar{w}\right)_{L_2(\Omega)} - A(\eta)\xi\right\} = \frac{1}{t}\left\{\frac{D}{h}\,t\,(\Delta\bar{w},\Delta\tilde{w})_{L_2(\Omega)} -\right.$$

$$-\lambda\,t\,\left([F_0,\tilde{w}],\bar{w}\right)_{L_2(\Omega)} - t\,\left([k,\tilde{w}],\bar{F}\right)_{L_2(\Omega)} -$$

$$-t\,\left([\bar{F},\tilde{w}],w\right)_{L_2(\Omega)} - t\left([F,\tilde{w}],\bar{w}\right)_{L_2(\Omega)} - t^2\,\left([\bar{F},\tilde{w}],\bar{w}\right)_{L_2(\Omega)} -$$

$$-\frac{1}{E}\,t\,(\Delta\bar{F},\Delta\tilde{F})_{L_2(\Omega)} - t\,\left([k,\bar{F}],\tilde{w}\right)_{L_2(\Omega)} - t\,\left([w,\bar{w}],\tilde{F}\right)_{L_2(\Omega)} -$$

$$\left.-\frac{1}{2}\,t^2\,\left([\bar{w},\tilde{w}],\tilde{F}\right)_{L_2(\Omega)}\right\}.$$

Wie gleich zu sehen ist, existiert der Limit dieses Quotienten für $t \to 0$ und es gilt

$$V\,A(\eta,\xi)\xi = \frac{D}{h}(\Delta\bar{w},\Delta\tilde{w})_{L_2(\Omega)} - \lambda\left([F_0,\bar{w}],\tilde{w}\right)_{L_2(\Omega)} -$$

$$-\left([k,\bar{F}],\tilde{w}\right)_{L_2(\Omega)} - \left([\bar{F},w],\tilde{w}\right)_{L_2(\Omega)} - \left([F,\bar{w}],\tilde{w}\right)_{L_2(\Omega)} -$$

$$-\frac{1}{E}(\Delta\bar{F},\Delta\tilde{F})_{L_2(\Omega)} - \left([k,\bar{w}],\tilde{F}\right)_{L_2(\Omega)} - \left([w,\bar{w}],\tilde{F}\right)_{L_2(\Omega)}.$$

Es ist ersichtlich, dass $V\,A(\eta,\xi)\xi$ in Bezug auf ξ linear ist, also $V\,A(\eta,\xi)\xi = D\,A(\eta,\xi)\xi$. Durch Anwendung der Hölderischen Ungleichlung erhalten wir die Stetigkeit mit Rücksicht auf η. Die Gleichheit $D\,A(\eta,\xi)\xi = D\,A(\eta,\xi)\xi$ zeigen wir durch Berechnung.

$$D\,A(\eta,\xi)\,\xi = \frac{D}{h}\left(\Delta\tilde{w},\Delta\bar{w}\right)_{L_2(\Omega)} - \lambda\left([F_0,\tilde{w}],\bar{w}\right)_{L_2(\Omega)} +$$

$$+ \left([k,\tilde{F}],\tilde{w}\right)_{L_2(\Omega)} - \left([\tilde{F},w],\bar{w}\right)_{L_2(\Omega)} - \left([F,\tilde{w}],\bar{w}\right)_{L_2(\Omega)} -$$

$$- \frac{1}{E}\left(\Delta\tilde{F},\Delta\bar{F}\right)_{L_2(\Omega)} - \left([k,\tilde{w}],\bar{F}\right)_{L_2(\Omega)} - \left([w,\tilde{w}],\bar{F}\right)_{L_2(\Omega)} =$$

$$= \frac{D}{h}\left(\Delta\bar{w},\Delta\tilde{w}\right)_{L_2(\Omega)} - \lambda\left([F_0,\bar{w}]\,\tilde{w}\right)_{L_2(\Omega)} - \left([k,\bar{F}],\tilde{w}\right)_{L_2(\Omega)} -$$

$$- \left([\bar{F},w],\tilde{w}\right)_{L_2(\Omega)} - \left([F,\bar{w}],\tilde{w}\right)_{L_2(\Omega)} - \frac{1}{E}\left(\Delta\bar{F},\Delta\tilde{F}\right)_{L_2(\Omega)} -$$

$$- \left([k,\bar{w}],\tilde{F}\right)_{L_2(\Omega)} - \left([w,\bar{w}],\tilde{F}\right)_{L_2(\Omega)} = D\,A(\eta,\xi)\,\xi\,.$$

In unserem Falle können wir das Funktional $\phi(\eta)$ leicht finden, es wird

$$\phi(w,F) = \frac{D}{2h}\left(\Delta w\right)^2_{L_2(\Omega)} - \frac{1}{2E}\left(\Delta F\right)^2_{L_2(\Omega)} - \frac{\lambda}{2}\left([F_0,w]\,;\right.$$

$$\left.,w\right)_{L_2(\Omega)} - \frac{1}{2}\left([k,F],w\right)_{L_2(\Omega)} - \frac{1}{3}\left([F,w],w\right)_{L_2(\Omega)} -$$

$$- \frac{1}{6}\left([w,w],F\right)_{L_2(\Omega)} - \frac{1}{2}\left([k,w],F\right)_{L_2(\Omega)}$$

sein.

Durch Verwendung des Hilfssatzes 2 können wir das Funktional $\phi(\eta)$ in die Form

$$\phi(w,F) = \frac{D}{2h}\left(\Delta w\right)^2_{L_2(\Omega)} - \frac{1}{2E}\left(\Delta F\right)^2_{L_2(\Omega)} - \frac{\lambda}{2}\left([F_0,w]\,,w\right)_{L_2(\Omega)} -$$

$$- \frac{1}{2}\left([F,w],w\right)_{L_2(\Omega)} - \left([k,F],w\right)_{L_2(\Omega)}\,, \qquad /20/$$

überschreiben.

6. Eigenschaften des Funktionals /20/

Das Funktional $\phi(\eta)$ können wir auch in folgender Form schreiben:

$$\phi(w,F) = \frac{D}{2h}(\Delta w)^2_{L_2(\Omega)} - \frac{1}{2E}(\Delta F)^2_{L_2(\Omega)} + \frac{\lambda}{2}\left\{\int_\Omega \left(\frac{\partial^2 F_0}{\partial x^2}\frac{\partial w}{\partial y} - \right.\right.$$

$$\left. - \frac{\partial^2 F_0}{\partial x \partial y}\frac{\partial w}{\partial x}\right)\frac{\partial w}{\partial y}\,d\Omega + \int_\Omega \left(\frac{\partial^2 F_0}{\partial y^2}\frac{\partial w}{\partial x} - \frac{\partial^2 F_0}{\partial x \partial y}\frac{\partial w}{\partial y}\right)\frac{\partial w}{\partial x}\,d\Omega\Big\} +$$

$$+ \frac{1}{2}\left\{\int_\Omega \left[\left(\frac{\partial^2 F}{\partial x^2}\frac{\partial w}{\partial y} - \frac{\partial^2 F}{\partial x \partial y}\frac{\partial w}{\partial x}\right)\frac{\partial w}{\partial y}\,d\Omega + \right.\right. \qquad /21/$$

$$\left.\left. + \int_\Omega \left(\frac{\partial^2 F}{\partial y^2}\frac{\partial w}{\partial x} - \frac{\partial^2 F}{\partial x \partial y}\frac{\partial w}{\partial y}\right)\frac{\partial w}{\partial x}\,d\Omega\right\} - \left([k,F],w\right)_{L_2(\Omega)}$$

Definition 6. Es sei im reellen Banachschen Raum B das Funktional ϕ gegeben. Wir sagen, dass ϕ in $\eta \in B$ von unten schwach halbstetig ist, wenn für jede Folge $\{\eta_n\}_1^\infty \subset B$ die Implikation

$$\left(\eta_n \xrightarrow[n\to\infty]{} \eta\right) \Longrightarrow \left(\phi(\eta) \leq \lim_{n\to\infty} \inf \phi(\eta_n)\right).$$

gilt.

Hilfssatz 3. ([19], § 18). Das Funktional ϕ habe in irgendeiner Kugel $D_r \subset B$ das erste und zweite Gateaux-sche Differential. Das zweite Differential des Funktionals ϕ sei in D_r nichtnegativ. Dann ist ϕ auf D_r schwach halbstetig von unten.

Satz 2. Das Funktional $\phi(w,F)$ ist auf W schwach halbstetig von unten.

Beweis. Bezeichnen wir das Funktional /21/ in der Form

$$\phi(w,F) = \sum_{i=1}^{5} I_i, \qquad /22/$$

wo $\hat{I}_i$ (i = 1,2,...5) die einzelnen Summanden sind. Wir sehen dann gleich, dass die Funktionale I_1, I_2 in $W = (H)^2$ gemäss Hilfssatz 3 schwach halbstetig von unten sind. Es sei $F_n \xrightarrow[n\to\infty]{} F$ in H. Also auch in $L_2(\Omega)$. Gemäss den bekannten Sätzen über die schwache Konvergenz [18] gilt dann

$$\forall_n \left(\|F_n\|_{W_2^{(2)}(\Omega)} \le C_F \right), \qquad \|F\|_{W_2^{(2)}(\Omega)} \le C_F,$$

C_F ist eine Konstante.

Dann ist also

$$\left\| \frac{\partial^2 F}{\partial x^2} \right\|_{L_2(\Omega)} \le C_{1F} , \quad \forall_n \left(\left\| \frac{\partial^2 F_n}{\partial x^2} \right\|_{L_2(\Omega)} \le C_{1F} \right),$$

C_{1F} ist eine Konstante.

$L_2(\Omega)$ ist ein reflexiver Raum, deshalb ist es möglich aus jeder begrenzten Folge in demselben eine schwach konvergente auszuwählen. Aus der Möglichkeit der Wahl einer schwach konvergenten Folge und aus der Definition der verallgemeinerten Ableitung $\frac{\partial^2 F}{\partial x^2}$ folgt, dass die ganze Folge $\left\{ \frac{\partial^2 F_n}{\partial x^2} \right\}_1^\infty$ in $L_2(\Omega)$ schwach konvergent ist und in $L_2(\Omega)$ den Limit $\frac{\partial^2 F}{\partial x^2}$ hat.

Ganz ähnlich können wir beweisen, dass auch

$$\frac{\partial^2 F_n}{\partial x \partial y} \xrightarrow[L_2(\Omega)]{} \frac{\partial^2 F}{\partial x \partial y} , \quad \frac{\partial^2 F_n}{\partial y^2} \xrightarrow[L_2(\Omega)]{} \frac{\partial^2 F}{\partial y^2}$$

ist. Es sei $w_n \longrightarrow w$ in H. Den Raum H können wir stetig in $\overset{o}{W}{}_2^{(1)}(\Omega)$ einbetten. Den Raum $\overset{o}{W}{}_2^{(1)}(\Omega)$ können wir kompakt in $W_4^{(1)}(\Omega)$ einbetten. Bei einer kompakten Abbilgung geht die schwach konvergente Folge in eine stark konvergente über, es ist also $w_n \longrightarrow w$ in $W_4^{(1)}(\Omega)$. Daraus also haben wir, dass

$$\frac{\partial w_n}{\partial x} \xrightarrow[L_4(\Omega)]{} \frac{\partial w}{\partial x}$$

Vollkommen ähnlich wäre der Beweis für $\frac{\partial w_n}{\partial y}$.

Es gilt folgende behannte Implikation

$$\left(\eta_n \xrightarrow[L_2(\Omega)]{} \eta \,,\ \xi_n \xrightarrow[L_2(\Omega)]{} \xi \Rightarrow \left(\eta_n\, \xi_n \xrightarrow[L_2(\Omega)]{} \eta\, \xi\right)\right).$$

Durch Verwendung dieser Implikation und aus der Definition der schwachen Konvergenz in $L_2(\Omega)$ erhalten wir sogleich, dass die Funktionale I_3, I_4, I_5 aus /22/ auf W schwach stetig sind. Also das Funktional /21/ ist auf W schwach halbstetig von unten.

F o l g e r u n g 1. Das Funktional /21/ ist auf $W = H \times H$ stetig.

Den Begriff der schwachen Lösung des Systems /1/, /2/, bei den Randbedingungen /3/ können wir folgend einführen:

D e f i n i t i o n 7. Das Paar $\langle w,F\rangle \in W$ werden wir eine schwache Lösung der Aufgabe /1/. /2/, /3/ nennen, wenn für jedes Paar $\langle \tilde{w}, \tilde{F}\rangle \in W$

$$\frac{D}{h}\int_\Omega \Delta w \Delta\tilde{w}\, d\Omega = \lambda\int_\Omega [F_0,w]\tilde{w}\, d\Omega + \int_\Omega [k,F]\tilde{w}\, d\Omega + \int_\Omega [F,w]\tilde{w}\, d\Omega, \qquad /23/$$

$$-\frac{1}{E}\int_\Omega \Delta F\Delta\tilde{F}\, d\Omega = \frac{1}{2}\int_\Omega [w,w]\,\tilde{F}\, d\Omega + \int_\Omega [k,w]\,\tilde{F}\, d\Omega$$

gilt.

Die erste Gleichung des Systems /23/ werden wir in der Form

$$[w,\tilde{w}]_\Delta = \sum_{i=1}^{3}\left(\lambda J_i + K_i + L_i\right) \qquad /24/$$

schreiben, wo

$$J_1 = \int_\Omega \frac{\partial^2 F_o}{\partial x^2} \frac{\partial^2 w}{\partial y^2} \tilde{w}\, d\Omega ,$$

$$J_2 = \int_\Omega \frac{\partial^2 F_o}{\partial y^2} \frac{\partial^2 w}{\partial x^2} \tilde{w}\, d\Omega ,$$

$$J_3 = -2\int_\Omega \frac{\partial^2 F_o}{\partial x \partial y} \frac{\partial^2 w}{\partial x \partial y} \tilde{w}\, d\Omega ,$$

$$K_1 = \int_\Omega \frac{\partial^2 k}{\partial x^2} \frac{\partial^2 F}{\partial y^2} \tilde{w}\, d\Omega ,$$

$$K_2 = \int_\Omega \frac{\partial^2 k}{\partial y^2} \frac{\partial^2 F}{\partial x^2} \tilde{w}\, d\Omega ,$$

$$K_3 = -2\int_\Omega \frac{\partial^2 k}{\partial x \partial y} \frac{\partial^2 F}{\partial x \partial y} \tilde{w}\, d\Omega ,$$

$$L_1 = \int_\Omega \frac{\partial^2 F}{\partial x^2} \frac{\partial^2 w}{\partial y^2} \tilde{w}\, d\Omega ,$$

$$L_2 = \int_\Omega \frac{\partial^2 F}{\partial y^2} \frac{\partial^2 w}{\partial x^2} \tilde{w}\, d\Omega ,$$

$$L_3 = -2\int_\Omega \frac{\partial^2 F}{\partial x \partial y} \frac{\partial^2 w}{\partial x \partial y} \tilde{w}\, d\Omega .$$

Nehmen wir das Funktional J_1. Durch Verwendung des Satzes über die Einbettung des Raumes $W_2^{(2)}(\Omega)$ in $C(\Omega)$ und der Cauchy-Schwarz-schen Ungleichheit erhalten wir

$$|J_1(w,\tilde{w})| \le \max_{\bar{\Omega}} |\tilde{w}| \left| \int_\Omega \frac{\partial^2 F_o}{\partial x^2} \frac{\partial^2 w}{\partial y^2} d\Omega \right| \le \qquad /25/$$

$$\le c_1 \max_{\bar{\Omega}} |\tilde{w}| \left| \int_\Omega \frac{\partial^2 F_o}{\partial x^2} d\Omega \right| \|w\|_o \le c_2 \|w\|_o \|\tilde{w}\|_o ,$$

/wir setzen voraus, dass $F_0(x,y)$ eine hinreichend reguläre Funktion ist/. Aus diesen Ungleichheiten ist zu sehen, dass das Funktional $J_1(w,\tilde{w})$ für ein festes $w \in H$ begrenzt ist. Aus der Form $J_1(w,\tilde{w})$ ist auch ersichtlich, dass es mit Rücksicht auf beide Veränderliche linear ist. Aus dem Reisz-schen Repräsentationssatz erhalten wir, dass ein einziges Element $P_1 w \in H$ derart existiert, dass

$$\left[P_1 w,\tilde{w}\right]_\Delta = J_1(w,\tilde{w}) \qquad /26/$$

für alle $\tilde{w} \in H$ gilt. Für den Operator P_1 gilt

$$\|P_1 w\|_0 \leq C_2 \|w\|_0 , \qquad w \in H. \qquad /27/$$

Also ist $P_1 \in (H \longrightarrow H)$ ein begrenzter Operator und mit Rücksicht auf die Linearität des Funktionals $J_1(w , \tilde{w})$ von w ist er linear und also stetig.

Durch ein ähnliches Verfahren für die Funktionale $J_2(w,\tilde{w})$, $J_3(w,\tilde{w})$ könnten wir linear stetige Operatoren P_2, $P_3 \in (H \rightarrow H)$ erhalten.

Bezeichnen wir $P = \sum_{i=1}^{3} P_i$

Satz 3. Der Operator P ist kompakt.

Beweis. Der Beweis wird nur für P_1 durchgeführt. Für die übrigen P_i $(i = 2,3)$ wäre er ähnlich. Es sei $\{w_n\}_1^\infty \subset H$ und es sei $\forall_n \|w_n\|_0 \leq C_H$ (eine feste Konstante.)

Aus /27/ haben wir dann

$$\forall_n \|P_1 w_n\|_0 \leq C_2 C_H .$$

Da der Einbettungsoperator von $W_2^{(2)}(\Omega)$ bis $L_2(\Omega)$ kompakt ist, erhalten wir, dass wir aus der Folge $\{P_1 w_n\}_1^\infty$ eine Konvergente in der Norm L_2 auswählen können.

Bezeichnen wir sie wieder $\left\{P_1 w_n\right\}_1^{\infty}$. Durch Verwendung der Cauchy-Schwarz-schen Ungleichheit erhalten wir dann

$$\|P_1 w_m - P_1 w_n\|_o^2 \leq C_3 \|w_m - w_n\|_o \|P_1 w_m - P_1 w_n\|_{L_2(\Omega)} \leq 2C_3C_4$$

$$\|P_1 w_m - P_1 w_n\|_{L_2(\Omega)} .$$

Aus den letzten Ungleichheiten ist ersichtlich, dass $\left\{P_1 w_n\right\}_1^{\infty}$ in H konvergent und also P_1 kompakt ist.

Folgerung 2. P ist vollstetig.

Nehmen wir jetzt das Funktional $K_1(F, \tilde{w})$. Auf die gleiche Weise wie für das Funktional $J_1(w,\tilde{w})$ können wir uns überzeugen, dass folgende Abschätzungen gelten

$$|K_1(F,\tilde{w})| \leq \max_{\Omega} |\tilde{w}| \left| \int_{\Omega} \frac{\partial^2 k}{\partial x^2} \frac{\partial^2 F}{\partial y^2} \, d\Omega \right| \leq$$

$$\leq C_4 \max_{\Omega} |\tilde{w}| \left| \int_{\Omega} \frac{\partial^2 k}{\partial x^2} \, d\Omega \right| \|F\|_o . \qquad /28/$$

Über die Krümmung k (x,y) setzen wir voraus, dass sie eine hinreichend reguläre Funktion ist, deshalb erhalten wir aus den letzten Ungleichheiten und durch Verwendung von /16/

$$|K_1(F,\tilde{w})| \leq C_5 \|w\|_o^2 \|\tilde{w}\|_o . \qquad /29/$$

Bei festem $w \in H$ ist also das Funktional $K_1(F,w)$ linear mit Rücksicht auf $\tilde{w}$ und begrenzt. Durch Verwendung des Riesz-schen Satzes erfahren wir, dass gerade ein solches Element $Q_1 w \in H$ derart existiert, dass

$$K_1(F,\tilde{w}) = [Q_1 w, \tilde{w}]_{\Delta} . \qquad /30/$$

Der gewonnene Operator Q_1 ist ersichtlich nichtlinear, operiert von H zu H (wir werden $Q_1 \in (H \to H)$ bezeichnen) und aus /29/ und /30/ folgt die Ungleichheit

$$\| Q_1 w \|_o \leq c_5 \| w \|_o^2 . \qquad /31/$$

Durch ein ähnliches Verfahren für die Funktionale $K_2(F,\tilde{w})$, $K_3(F,\tilde{w})$ würden wir die Operatoren Q_2 Q_3 erhalten, für welche ähnliche Abschätzungen wie /31/ gelten.

Bezeichnen wir

$$Q = \sum_{i=1}^{3} Q_i$$

Satz 4. Der Operator Q ist kompakt.

B e w e i s. Es genügt die Kompaktheit des Operators Q_1 zu beweisen, für die übrigen Q_i (i = 2,3) wäre der Beweis ähnlich. Es sei $\{w_n\}_1^\infty \subset H$ und es sei $\forall_n \| w_n \|_o \leq c_H$. Aus /31/ erhalten wir sogleich

$$\forall_n \| Q_1 w_n \|_o \leq c_5 c_H^2 .$$

Aus der letzten Ungleichheit und aus der totalen Stetigkeit des Einbettungsoperators von $W_2^{(2)}(\Omega)$ bis $C(\Omega)$ erhalten wir die Kompaktheit der Folge $\{Q_1 w_n\}_1^\infty$ in $C(\Omega)$. Es existiert also eine solche aus der Folge $\{Q_1 w_n\}_1^\infty$ ausgewählte Unterfolge (wir werden sie wieder $\{Q_1 w_n\}_1^\infty$ bezeichnen) dass sie in $C(\Omega)$ Cauchysch ist. Dann aber erhalten wir aus /30/,/28/und /29/ die Abschätzungen

$$\| Q_1 w_m - Q_1 w_n \|_o^2 \leq \max_{\bar{\Omega}} | Q_1 w_m - Q_1 w_n | c_H^2 . \qquad /32/$$

Aus /32/ ist ersichtlich, dass $\{Q_1 w_n\}_1^\infty$ in H Cauchysch ist, Q_1 ist also kompakt.

Satz 5. Der Operator Q ist stetig.

Beweis. Wieder genügt es nur die Stetigkeit des Operators Q_1 zu beweisen.

Es sei ein beliebiges Element $w \in H$ gegeben und die Folge $\{w_n\}_1^\infty \subset H$ sei derart, dass $\|w_n - w\|_0 \xrightarrow[n\to\infty]{} 0$. Für ein beliebiges $\tilde{w} \in H$ gilt

$$\left|[Q_1 w_n - Q_1 w, \tilde{w}]_\Delta\right| = \left|K_1(F(w_n),\tilde{w}) - K_1(F(w),\tilde{w})\right| =$$

$$= \left|\int_\Omega \frac{\partial^2 k}{\partial x^2}\frac{\partial^2}{\partial y^2}\left(F(w_n) - F(w)\right)\tilde{w}\,d\Omega\right| \le \max_{\bar{\Omega}}|\tilde{w}|\left|\int_\Omega \frac{\partial^2 k}{\partial x^2}\frac{\partial^2}{\partial y^2}\left(F(w_n) - F(w)\right)d\Omega\right| \le C_4 \max_{\bar{\Omega}}|\tilde{w}|\left|\int_\Omega \frac{\partial^2 k}{\partial x^2}\,d\Omega\right|\|F(w_n) - F(w)\|_0 \le$$

$$\le C_7 \|\tilde{w}\|_0 \|F(w_n) - F(w)\|_0 .$$

Daraus aber erhalten wir gleich, dass

$$\|Q_1 w_n - Q_1 w\|_0 \le C_7 \|F(w_n) - F(w)\|_0 .$$

Durch Verwendung der Beziehungen /14/, /15/ erhalten wir

$$\|F(w_n) - F(w)\|_0 \le E\alpha\left\{\int_\Omega\left|\left(\frac{\partial^2 w}{\partial x\partial y} - \frac{\partial^2 w_n}{\partial x\partial y}\right)\left(\frac{\partial^2 w}{\partial x\partial y} - \frac{\partial^2 w_n}{\partial x\partial y}\right) + \left(\frac{\partial^2 w_n}{\partial x^2} - \frac{\partial^2 w}{\partial x^2}\right)\frac{\partial^2 w_n}{\partial y^2} - \left(\frac{\partial^2 w_n}{\partial y^2} - \frac{\partial^2 w}{\partial y^2}\right)\frac{\partial^2 w}{\partial x^2}\,d\Omega\right| + A\|w_n - w\|_0\right\} \le E\alpha\left\{\int_\Omega\left|\left(\frac{\partial^2 w}{\partial x\partial y}\right)^2 - \left(\frac{\partial^2 w_n}{\partial x\partial y}\right)^2\right|d\Omega + \int_\Omega\left|\left(\frac{\partial^2 w_n}{\partial x^2} - \frac{\partial^2 w}{\partial x^2}\right)\frac{\partial^2 w_n}{\partial y^2} - \left(\frac{\partial^2 w_n}{\partial y^2} - \frac{\partial^2 w}{\partial y^2}\right)\frac{\partial^2 w}{\partial x^2}\right|d\Omega + A\|w_n - w\|_0\right\} \le \|w_n - w\|_0 E\alpha\left\{2\varkappa^2\left(\|w_n\|_0 + \|w\|_0\right) + A\right\} \qquad /33/$$

Also gilt die Ungleichheit

$$\|Q_1 w_n - Q_1 w\|_0 \leq c_7 E\alpha \left\{ 2\mathcal{H}^2 \left(\|w\|_0 + \|w_n\|_0 \right) + A \right\} \|w_n - w\|_0,$$

welche den Satz 5 beweist.

F o l g e r u n g 3. Der Operator Q ist vollstetig

Durch Verwendung des Satzes über die Einbettung und der Caucha--Schwarzschen Ungleichheit können folgende Abschätzungen gemacht werden:

$$\left| L_1 (w,\tilde{w}) \right| \leq \max_{\bar{\Omega}} |\tilde{w}| \left| \int_\Omega \frac{\partial^2 F}{\partial x^2} \frac{\partial^2 w}{\partial y^2} \, d\Omega \right| \leq c_7 \max_{\bar{\Omega}} |\tilde{w}| \, \|F\|_0 \, \|w\|_0 .$$

Daraus erhalten wir ähnlich wie in /29/

$$\left| L_1(w,\tilde{w}) \right| \leq c_8 \|w\|_0^3 \|\tilde{w}\|_0. \qquad /34/$$

Bei festem $w \in H$ ist $L_1(w,\tilde{w})$ in Bezug auf w ein lineares und begrentztes Funktional. Gemäss dem Rieszschen Satz existiert gerade ein Element $R_1 w \in H$ derart, dass

$$L_1(w,\tilde{w}) = \left[R_1 w, \tilde{w} \right]_\Delta . \qquad /35/$$

Der Operator R_1 ist nichtlinear und $R_1 \in (H \longrightarrow H)$. Aus /34/ und /35/ folgt

$$\|R_1 w\|_0 \leq c_8 \|w\|_0^3 . \qquad /36/$$

Durch Wiederholung desselben Vorganges für $L_2 (w,\tilde{w})$, $L_3(w,\tilde{w})$ erhalten wir leicht die Operatoren R_2 , R_3, welche abermals nichtlinear sein werden und für welche ähnliche Abschätzungen wie /36/ gelten.

Wir bezeichnen $R = R_1 + R_2 + R_3$.

S a t z 6. Der Operator R ist kompakt.

B e w e i s . Dieser ist dem Beweis des Satzes 4 vollkommen ähnlich.

S a t z 7. Der Operator R ist stetig.

B e w e i s. Wir werden nur die Stetigkeit von R_1 beweisen. Es sei $w \in H$ /ein beliebiges aber festes Element/, $\{w_n\}_1^\infty \subset H$ und es sei $\|w_n - w\|_0 \xrightarrow[n\to\infty]{} 0$. Für ein beliebiges $\tilde{w} \in H$ gilt

$$\left|[R_1 w_n - R_1 w, \tilde{w}]_\Delta\right| = \left|\int_\Omega \left(\frac{\partial^2 F(w_n)}{\partial x^2}\frac{\partial^2 w_n}{\partial y^2} - \frac{\partial^2 F(w)}{\partial x^2}\frac{\partial^2 w}{\partial y^2}\right)\tilde{w}\, d\Omega\right| \le$$

$$\le \max_{\bar{\Omega}} |\tilde{w}| \left|\int_\Omega \left[\left(\frac{\partial^2 F(w_n)}{\partial x^2} - \frac{\partial^2 F(w)}{\partial x^2}\right)\frac{\partial^2 w_n}{\partial y^2} + \left(\frac{\partial^2 w_n}{\partial y^2} -\right.\right.\right.$$

$$\left.\left.\left. - \frac{\partial^2 w}{\partial y^2}\right)\frac{\partial^2 F(w)}{\partial x^2}\right] d\Omega\right| \le c_9 \|\tilde{w}\|_0 \left\{\left|\int_\Omega \left[\frac{\partial^2}{\partial x^2}\left(F(w_n) - F(w)\right)\frac{\partial^2 w_n}{\partial y^2}\right] d\Omega\right| +\right.$$

$$\left. + \left|\int_\Omega \left[\frac{\partial^2}{\partial y^2}(w_n - w)\frac{\partial^2 F(w)}{\partial x^2}\right] d\Omega\right|\right\} \le c_{10}\|\tilde{w}\|_0 \left[\|F(w_n) -\right.$$

$$\left. - F(w)\|_0 \|w_n\|_0 + \|w_n - w\|_0 \|F(w)\|_0\right].$$

Wenn wir $\tilde{w} = R_1 w_n - R_1 w$ setzen, erhalten wir

$$\|R_1 w_n - R_1 w\|_0^2 \le c_{10}\|R_1 w_n - R_1 w\|_0 \left\{\|F(w_n) - F(w)\|_0 \|w_n\|_0 +\right.$$

$$\left. + \|w_n - w\|_0 \|F(w)\|_0\right\}.$$

Durch Verwertung des Vorganges /33/, durch Kürzen und Verwertung der Beziehung /16/ geht die letzte Ungleichheit in die Form

$$\|R_1 w_n - R_1 w\|_0 \le c_{10}\left\{E\alpha\left[2\mathcal{K}^2\left(\|w_n\|_0 + \|w\|_0\right) + A\right]\|w_n\|_0 +\right.$$

$$\left. + 2E\mathcal{K}^2\alpha\left(\|w\|_0 + A\right)\|w\|_0\right\}\|w_n - w\|_0$$

über, womit der Satz bewiesen ist.

F o l g e r u n g 4. Der Operator R ist vollstetig. Wenn wir die Ergebnisse der vorhergehenden Sätze vereinige, erhalten wir

S a t z 8. Die erste Gleichung des Systems /23/ können wir in H in der äquivalenten Operatorform

$$w = T w = \lambda P w + Q w + R w$$

schreiben, wo T ein vollstetiger Operator ist.

B e w e i s. Nach dem vorhergegangenen Verfahren können wir die Gleichung /24/ in H in der Form

$$[w,\tilde{w}]_\Delta = \lambda [P w,\tilde{w}]_\Delta + [Q w,\tilde{w}]_\Delta + [R w,\tilde{w}]_\Delta \qquad /37/$$

schreiben. Aber die Gleichung /37/ ist in H äquivalent mit der Operatorgleichung

$$w - \lambda P w - Qw - Rw = 0, \quad w \in H.$$

Die Eigenschaft des Operators T geht aus den Folgerungen 2,3,4 hervor.

Es sei ϕ_1 ein Funktional auf dem Banachraum B.

H i l f s s t z 4 ([19], § 8). Aus der Kompaktheit grad ϕ_1 in der Kugel $Dr \in B$ folgt, dass ϕ_1 in dieser Kugel schwach stetig ist.

Es seien $F_1(x,y)$, $F_2(x,y)$ solche Funktionen, dass bei den gegebenen Randbedingungen aus /3/

$$\frac{1}{E} \Delta\Delta F_1 = -\frac{1}{2} [w,w] , \frac{1}{E} \Delta\Delta F_2 = - [k,w] , w \in H \qquad /38/$$

gilt. Es kann gezeigt werden /z.B. aus der Teorie der linearen Randwertprobleme mit positiv definiten Operator [20]) , dass bei so regulären rechten Seiten, wie sie die Gleichungen in /38/ besitzen, die Funktionen $F_1(x,y)$, $F_2(x,y)$ immer existieren und eindeutig gegeben sind. Das Funktional /20/ können wir also in der Form

$$\phi(w) = \frac{D}{2h}(\Delta w)^2_{L_2(\Omega)} - \frac{1}{2E}(\Delta F_1{}^2)_{L_2(\Omega)} - \frac{1}{2E}(\Delta F_2)^2_{L_2(\Omega)} +$$

$$+ \frac{1}{4E}([w,w], F_2)_{L_2(\Omega)} + \frac{1}{2E}([k,w], F_1)_{L_2(\Omega)} - \qquad /39/$$

$$- \frac{1}{2}([F_1,w], w)_{L_2(\Omega)} - \frac{1}{2}([F_2,w], w)_{L_2(\Omega)} -$$

$$- \frac{\lambda}{2}([F_o,w], w)_{L_2(\Omega)} - ([k,w], F_1 + F_2)_{L_2(\Omega)}$$

schreiben. Aus dem Hilfssatz 4 und dem Satz 8 erhalten wir dann die

F o l g e r u n g 5. Das Funktional $\phi_1(w) = \phi(w) - \frac{D}{2h}(w)^2_{L_2(\Omega)}$ ist in H schwach stetig. Das Funktional /39/ können wir in die Form

$$\phi(w) = \frac{D}{2h}(\Delta w)^2_{L_2(\Omega)} - \frac{1}{2E}(\Delta F_1, \Delta F_1)_{L_2(\Omega)} - \frac{1}{2E}(\Delta F_2, \Delta F_2)_{L_2(\Omega)} -$$

$$- \frac{1}{2E}(\Delta F_1, \Delta F_2)_{L_2(\Omega)} + \frac{1}{2E}(\Delta F_2, \Delta F_1)_{L_2(\Omega)} + \frac{1}{2E}(\Delta F_1, \Delta F_1)_{L_2(\Omega)} +$$

$$+ \frac{1}{E}(\Delta F_1, \Delta F_2)_{L_2(\Omega)} + \frac{1}{E}(\Delta F_2, \Delta F_2)_{L_2(\Omega)} - \frac{\lambda}{2}([F_o,w],w)_{L_2(\Omega)} =$$

$$= \frac{D}{2h}\|w\|^2_H + \frac{1}{2E}\left(\|F_1\|^2_H + \|F_2\|^2_H\right) - \frac{1}{4}([w,F_2],w)_{L_2(\Omega)} -$$

$$- \frac{1}{2}([k,F_1],w)_{L_2(\Omega)} - \frac{\lambda}{2}([F_o,w],w)_{L_2(\Omega)}$$

geben. Also können wir das Funktional /39/ z.B. in der Form

$$\phi(w) = \sum_{i=1}^{5} J_i(w) \qquad /40/$$

schreiben, wo

$$J_1 = \frac{D}{2h} \|w\|_H^2 ,$$

$$J_2(w) = \frac{1}{2E} \|F_2\|_H^2 ,$$

$$J_3(w) = \frac{1}{2E} \|F_1\|_H^2 , \qquad /41/$$

$$J_4(w) = -\frac{1}{4}\left\{(w,F_2),w\right)_{L_2(\Omega)} + 2\left([k,F_1],w\right)_{L_2(\Omega)} ,$$

$$J_5(w) = -\frac{\lambda}{2}\left([F_0,w],w\right)_{L_2(\Omega)} .$$

S a t z 9. Das Funktional /39/ ist auf H wachsend.

B e w e i s. Die Funktionale $J_i(w)$ $(i = 1,2,3,4,5)$ sind homogen. Das Funktional $J_4(w)$ in /41/ ist auf H schwach stetig.

Nehmen wir auf H eine Kugelfläche mit dem Mittelpunkt in θ_H $(\|\theta_H\|_H = 0)$ und mit einheitlichem Halbmesser und bezeichnen wir sie S_1. Es sei $Z_1 \subset S_1$ und Z_1 sei eine Menge solcher Punkte, wo

$$\frac{\lambda}{2}\left([F_0,w],w\right)_{L_2(\Omega)} > 0$$

gilt. Z_1 kann auch eine leere Menge sein. Bezeichnen wir die schwache Abgeschlossenheit Z_1 in H als $\tilde{Z}_1$. Aus der schwachen Stetigkeit $J_4(w)$ auf H erhalten wir auf $\tilde{Z}_1$ (und also auch auf Z_1) die Ungleichheit

$$|J_4(w)| \leq C_1, \qquad /42/$$

wo C_1 eine entsprechende positive Konstante ist.

Das Funktional $J_3(w)$ können wir auch in der Form

$$J_3(w) = -\frac{1}{4}\left([w,w], F_1\right)_{L_2(\Omega)}$$

schreiben. Also wenn $J_3(w) = 0$ ist, dann ist fast überall entweder $F_1 = \theta_H$, oder $[w,w] = \theta_H$.

Dann erhalten wir [9], dass $w \equiv \Theta_H$ ist. Daraus folgt aber, dass $J_3(w) \geq 0$ und die Gleichheit kann nur wenn $w \equiv \Theta_H$ ist entstehen.

Hilfssatz 5. $\Theta_H \notin \tilde{Z}_1$.

Beweis des Hilfssatzes. Die Bedingung für die Menge Z_1 können wir auch in der Form

$$\|w\|_H^2 - \frac{\lambda}{2} \left([F_0, w], w\right)_{L_2(\Omega)} < 1$$

schreiben.

Es sollte $\Theta_H \in \tilde{Z}_1$ sein. Dann würde eine solche Folge $\{w_n\} \subset Z_1$ existieren, dass $w_n \xrightarrow[n \to \infty]{} \Theta_H$. Das Funktional $\left([F_0, w], w\right)_{L_2(\Omega)}$ ist gemäss Hilfssatz 4 schwach stetig auf H und $\left([F, \Theta_H], \Theta_H\right)_{L_2(\Omega)} = 0$. Also wäre für ein genügend grosses n $\|w_n\|_H < 1$, was mit der Definition Z_1 im Widerspruch steht.

Da $J_3(w)$ auf H auch schwach stetig ist, gelten in $\tilde{Z}_1$ /also auch in Z_1/ die Ungleichheiten

$$0 < C_2 \leq J_3(w) \leq C_3, \qquad /43/$$

wo C_2, C_3 entsprechende Konstanten sind.

Nehmen wir jetzt die Kugelfäche S_r, welche ihren Mittelpunkt in Θ_H hat und ihr Halbmesser $r > 0$. Mit Z_{1r} bezeichnen wir die Mittelprojektion der Menge Z_1 auf S_r. Mit Rücksicht darauf, dass die Funktionale /41/ homogen sind werden auf Z_{1r} die Ungleichheiten

$$\phi(w) \geq \sum_{i=1}^{3} J_i - |J_4| + J_5 \geq C_2 r^4 - C_1 r^3 = r^3(C_2 r - C_1) \qquad /44/$$

gelten. Aus /44/ erhalten wir, dass wenn $r > \frac{C_1}{C_2}$, dann gilt auf Z_{1r}

$$\phi(w) \geq C_4 r^3, \quad C_4 > 0. \qquad /45/$$

Als Komplement zu Z_1 auf S_1 bezeichnen wir Z_2. Auf Z_2 wird also

$$\frac{\lambda}{2}\left([F_0,w],w\right)_{L_2(\Omega)} \leq \frac{D}{2h}\left(\|w\|_H^2 - 1\right) \qquad /46/$$

gelten.

Bezeichnen wir mit Z_{2r} die Mittelprojektion Z_2 auf S_r.

Aus /46/ ist ersichtlich, dass auf Z_{2r} die Unglichung

$$\frac{D}{2h}\|w\|_H^2 - \frac{\lambda}{2}\left([F_0,w],w\right)_{L_2(\Omega)} \geq \frac{D}{2h} r^2 \qquad /47/$$

gelten wird. Aus ähnlichen Erwägungen für Z_2, wie die Erwägungen /42/, /43/ für Z_1 sind, bekommen wir die Abschätzungen auf Z_{2r} in der folgenden Form

$$\begin{aligned}\phi(w) &\geq \frac{D}{2h}\|w\|_H^2 + J_2 + J_3 - |J_4| - \\ &- \frac{\lambda}{2}\left([F_0,w],w\right)_{L_2(\Omega)} \geq \frac{D}{2h} r^2 + c_5 r^4 - c_6 r^3 \geq \\ &\geq r^3 (c_5 r - c_6), \end{aligned} \qquad /48/$$

wo c_5, c_6 entsprechende Konstanten sind. Also wenn $r > \frac{c_6}{c_5}$, dann

$$f(w) \geq c_7 r^3, \quad c_7 > 0. \qquad /49/$$

Die Ungleichheiten /45/, /49/ zeigen, dass wenn $r > \max\left(\frac{c_1}{c_2}, \frac{c_6}{c_5}\right)$ dann ist das Funktional /39/ auf H wachsend.

7. Probleme der linearisierten Aufgabe

Linearisieren wir das System der Gleichungen /1/, /2/ so, dass wir das Gateauxsche Differential des Systems im Punkte $\langle \Theta_H, \Theta_H \rangle \in W$ konstruieren. Wir erhalten die linearisierte Aufgabe der Form

$$\frac{D}{h} \Delta\Delta w = \lambda [F_0, w] + [k, F], \qquad /50/$$

$$-\frac{1}{E} \Delta\Delta F = [k, w]. \qquad /51/$$

In [14] war ein Satz formuliert worden, welcher in unserem Falle die folgende Form haben wird:

Satz 10. Der Operator der Gleichung /50/ sei in $L_2(\Omega)$ symmetrisch. Dann ist das Bifurkationspunktenproblem für die Aufgabe /1/, /2/, /3/ linearisierbar, jeder Eigenwert des Systems /50/, /51/ ist ein Punkt der Bifurkation für /1/, /2/, /3/ dem ein stetiger Zweig nichttrivialer Lösungen entspricht.

Definition 8. Den auf der linearen Menge $D_A \subset L_2(\Omega)$ definierten Linearoperator $A \; (A \in D_A \longrightarrow L_2(\Omega))$ nennen wir in $L_2(\Omega)$ symmetrisch, wenn

$$\forall_{u,v \in D_A} (A u, v)_{L_2(\Omega)} = (u, A v)_{L_2(\Omega)}$$

gilt.

Satz 11. Der Operator der Gleichung /50/ ist bei den Randbedingungen /3/ symmetrisch in $L_2(\Omega)$.

Beweis. Die Symmetrie des biharmonischen Operators war in [10] bewiesen worden. Die Symmetrie des Ausdrucks $[F_0, w]$ geht aus dem Hilfssatz 2 hervor. Es bleibt noch die Symmetrie des Summanden [k,F] zu beweisen. Aus dem Hilfssatz 2 gilt

$$\left([k,F(w_1)],w_2\right)_{L_2(\Omega)} = \left([k,w_2],F(w_1)\right)_{L_2(\Omega)} .$$

$$\left([k,w_2],F(w_1)\right)_{L_2(\Omega)} = -\frac{1}{E}\left(\Delta F(w_2),\Delta F(w_1)\right)_{L_2(\Omega)} =$$

$$= -\frac{1}{E}\left(\Delta F(w_1),\Delta F(w_2)\right)_{L_2(\Omega)} = \left(F(w_2),[k,w_1]\right)_{L_2(\Omega)} =$$

$$= \left(w_1,[k,F(w_2)]\right)_{L_2(\Omega)} ,$$

damit ist der Beweis beendet.

Ein ähnliches Ergebniss für die Gleichung des Typs a/ erhielt Vorovič [21]. In den Arbeiten [6,9] war für die Gleichungen schlanker Wände des Typs b/ gezeigt worden, dass links vom ersten Eigenwert der linearisierten Aufgabe keine nichttriviale Lösung der nichtlinearen Aufgabe existiert.

Untersuchen wir unseren Fall. Das Funktional /20/ ist auf W schwach halbstetig von unten. Das Funktional $\phi_1(w)$ aus der Folgerung 5 ist schwach stetig auf H und das Funktional /39/ ist auf H wachsend. Also hat /39/ für jedes reelle λ das minimum auf W,+/ welches die Lösung des Systems /1/, /2/ [15] ist. Es ist auch ersichtlich, dass /1/, /2/, /3/ immer eine triviale Lösung hat und auch $\phi(\Theta_H,\Theta_H)=0$.

Satz 12. Wenn dem ersten Eigenwert λ_1 der linearisierten Aufgabe /50/, /51/, /3/ eine solche Eigenfunktion φ_1 entspricht, dass

$$\frac{1}{2E}\left([\varphi_1,\varphi_1],F_2\right)_{L_2(\Omega)} + \frac{1-2E}{E}\left([k,\varphi_1],F_1\right)_{L_2(\Omega)} +$$

$$+ \left([F_2,\varphi_1],\varphi_1\right)_{L_2(\Omega)} \neq 0, \qquad \left([F_0,\varphi_1],\varphi_1\right)_{L_2(\Omega)} > 0,$$

dann existiert für $\lambda<\lambda_1$ eine nichttriviale Lösung der Aufgabe /1/, /2/, /3/.

+/ Ein ähnliches Ergebniss für eine ähnliche Aufgabe wurde auf andere Art und Weise in der Arbeit von M.S. Berger, Indiana Univ. Journ., Vol. 20, No 7, 1971 gezeigt. Der Verfasser erfuhr von diesem Ergebniss erst nach Einlieferung seiner Arbeit.

Folgerung 6. Das Stabilitätsproblem der Schalen ist nicht möglich durch die linearisation der nichtlinearen Ausgangsaufgabe zu lösen.

Beweis des Satzes. Es sei λ_1 der kleinste Eigenwert der linearisierten Aufgabe /50/, /51/, /3/. φ_1 sei eine diesem Eigenwert entsprechende Eigenfunktion. Setzen wir $w = C\,\varphi_1$ /C ist eine reelle Konstante/ und rechnen wir $\phi(C\,\varphi_1)$. Wir erhalten

$$\phi(C\,\varphi_1) = \frac{D}{2h} C^2 \,(\Delta\varphi_1)^2_{L_2(\Omega)} - \frac{1}{2E} C^2 \,(\Delta F_2)^2_{L_2(\Omega)} + \frac{1}{4E} C^3 \,([\varphi_1,$$

$$, \varphi_1], F_2)_{L_2(\Omega)} + \frac{1}{2E} C^3 \,([k, \varphi_1], F_1)_{L_2(\Omega)} - \frac{C^3}{2} \,([F_2, \varphi_1], \varphi_1)_{L_2(\Omega)} -$$

$$- \frac{\lambda}{2} C^2 \,([F_o, \varphi_1], \varphi_1)_{L_2(\Omega)} + C^4 \frac{1}{2E} (\Delta F_1)^2_{L_2(\Omega)} - \frac{C^4}{2} \,([F_1, \varphi_1], \varphi_1)_{L_2(\Omega)} -$$

$$- C^2 \,([k, \varphi_1], F_2)_{L_2(\Omega)} - C^3 \,([k, \varphi_1], F_1)_{L_2(\Omega)} .$$

Also

$$\phi(C\,\dot{\varphi}_1) = \frac{C^2}{2} \left\{ \frac{D}{h} (\Delta\varphi_1)^2_{L_2(\Omega)} - \frac{1}{E} (\Delta F_2)^2_{L_2(\Omega)} - \right.$$

$$\left. - 2\,([k, \varphi_1], F_2)_{L_2(\Omega)} - \lambda([F_o, \varphi_1], \varphi_1)_{L_2(\Omega)} \right\} + \frac{C^3}{2} \left\{ \frac{1}{2E} \,([\varphi_1,\right.$$

$$\left. , \varphi_1], F_2)_{L_2(\Omega)} + \frac{1-2E}{E} \,([k, \varphi_1], F_1)_{L_2(\Omega)} + ([F_2, \varphi_1], \varphi_1)_{L_2(\Omega)} \right\} +$$

$$+ \frac{C^4}{2} \left\{ \frac{1}{E} (\Delta F_1)^2_{L_2(\Omega)} - ([F_1, \varphi_1], \varphi_1)_{L_2(\Omega)} \right\} .$$

Suchen wir das Minimum $C^{-2}\,2\phi\,(C\,\varphi_1)$. Wenn wir das Extrem durch ein gewöhnliches Verfahren suchen erhalten wir

$$- \frac{1}{2E} ([\varphi_1, \varphi_1], F_2)_{L_2(\Omega)} + \frac{1-2E}{E} \,([k, \varphi_1], F_1)_{L_2(\Omega)} + ([F_2, \varphi_1], \varphi_1)_{L_2(\Omega)} =$$

$$= 2\,C \left\{ \frac{1}{E} (\Delta F_1)^2_{L_2(\Omega)} - ([F_1, \varphi_1], \varphi_1)_{L_2(\Omega)} \right\} .$$

Also haben wir

$$C = -\frac{\frac{1}{2E}\left([\varphi_1, \varphi_1], F_2\right)_{L_2(\Omega)} + \frac{1-2E}{E}\left([k, \varphi_1], F_1\right)_{L_2(\Omega)} + \left([F_2, \varphi_1], \varphi_1\right)_{L_2(\Omega)}}{2\left\{\frac{1}{E}\left(\Delta F_1\right)_{L_2(\Omega)} - \left([F_1, \varphi_1], \varphi_1\right)_{L_2(\Omega)}\right\}}$$

Bezeichnen wir

$$C = -\frac{M}{2N},$$

wo der Nenner stets verschieden von Null ist, weil φ_1 eine Eigenfunktion des Systems /50/, /51/ ist. Ausserdem ist aus /50/, /51/ ersichtlich, dass

$$\frac{D}{h}\left(\Delta\varphi_1\right)^2_{L_2(\Omega)} = \lambda_1\left([F_0, \varphi_1], \varphi_1\right)_{L_2(\Omega)} + \left([k, F_2], \varphi_1\right)_{L_2(\Omega)} =$$

$$= \lambda_1\left([F_0, \varphi_1], \varphi_1\right)_{L_2(\Omega)} + \left([k, \varphi_1], F_2\right)_{L_2(\Omega)}$$

gilt. Dadurch erhalten wir

$$\min \frac{2\,\phi(c\,\varphi_1)}{c^2} = (\lambda_1 - \lambda)\left([F_0, \varphi_1], \varphi_1\right)_{L_2(\Omega)} - \frac{M^2}{2N} + \frac{M^2}{4N}.$$

Wir können richtigstellen

$$N = \frac{1}{E}\left(\Delta F_1\right)^2_{L_2(\Omega)} - \left([F_1, \varphi_1], \varphi_1\right)_{L_2(\Omega)} = \frac{1}{E}\left(\Delta F_1\right)^2_{L_2(\Omega)} -$$

$$- \left([\varphi_1, \varphi_1], F_1\right)_{L_2(\Omega)} = \frac{1}{E}\left(\Delta F_1\right)^2_{L_2(\Omega)} + \frac{2}{E}\left(\Delta F_1\right)^2_{L_2(\Omega)} =$$

$$= \frac{3}{E}\left(\Delta F_1\right)^2_{L_2(\Omega)}.$$

Also

$$\min \frac{2\phi(c\,\varphi_1)}{c^2} = (\lambda_1 - \lambda)\,\left([F_o, \varphi_1], \varphi_1\right)_{L_2(\Omega)} - \frac{E\,M^2}{12\,(\Delta F_1)^2_{L_2(\Omega)}} .$$

Jetzt können folgende Fälle eintreten:

1. $\min\left(c^{-2}\,2\,\phi\,(c\,\varphi_1)\right) \geq 0 \Rightarrow$ kleinerer Wert $\phi\,(w)$ ist für $w = \Theta_H$. Dieser Fall jedoch besagt nichts, da die triviale Lösung die Lösung /1/. /2/. /3/ für alle λ ist.

2. $\min\left(c^{-2}\,2\,\phi\,(c\,\varphi_1)\right) < 0 \Rightarrow (\lambda_1 - \lambda)\left([F_o, \varphi_1], \varphi_1\right)_{L_2(\Omega)} < \frac{E}{12}\,\frac{M^2}{(\Delta F_1)^2_{L_2(\Omega)}}$.

2 a. $\left([F_o, \varphi_1], \varphi_1\right)_{L_2(\Omega)} < 0 \Rightarrow$

$$\Rightarrow \lambda_1 - \lambda > \frac{E}{12}\,\frac{M^2}{(\Delta F_1)^2_{L_2(\Omega)}}\,\frac{1}{\left([F_o, \varphi_1], \varphi_1\right)_{L_2(\Omega)}} \Rightarrow$$

$$\Rightarrow \lambda < \lambda_1 - \frac{E}{12}\,\frac{M^2}{(\Delta F_1)^2_{L_2(\Omega)}}\,\frac{1}{\left([F_o, \varphi_1], \varphi_1\right)_{L_2(\Omega)}} \Rightarrow$$

$\Rightarrow$ Wenn $M \neq 0$, dann

existiert eine nichttriviale Lösung rechts von λ_1. Dieser Fall wurde durch die linearisation der Aufgabe /1/, /2/, /3/ gelöst.

2 b. $\left([F_o, \varphi_1], \varphi_1\right)_{L_2(\Omega)} > 0$. Dann aber

$$\lambda > \lambda_1 - \frac{E}{12}\,\frac{M^2}{(\Delta F_1)^2_{L_2(\Omega)}} \cdot \frac{1}{\left([F_o, \varphi_1], \varphi_1\right)_{L_2(\Omega)}} .$$

Der Fall 2 b. besagt, dass wenn $M \neq 0$, für ein gewisses $\lambda < \lambda_1$ gilt, dass $\min \phi (C \varphi_1) < 0$. Dann muss aber auch das absolute Minimum $\phi (w)$ negativ sein. Also die Lösung der Aufgabe /1/, /2/, /3/ ist für $\lambda < \lambda_1$ nichttrivial /anderenfalls würde für die Lösung $w_0 = \Theta_H$, $\phi (w_0) = 0$ gelten/.

B e m e r k u n g 2. Numerische Experimente zeigen, dass im Falle der Gültigkeit des Satzes 12 die nichttrivialen Lösungen für $\lambda < \lambda_1$ einen stetigen Zweig mit wachsender Norm bilden.

L i t e r a t u r

1. VOĽMIR A.S., Ustojčivosť uprugich sistem, /russisch/, Moskva 1963.
2. VOROVIČ I.I., O susčestvovanii rešenij v nelinejnoj teorii oboloček, /russisch/, Izv.AN SSSR, ser. matem., 19, 1955.
3. VOROVIČ I.I., O nekotorych priamych metodach v nelinejnoj teorii pologich oboloček, /russisch/, Prikl.mat.i mech., XX, 4, 1956.
4. MOROZOV N.F., Nelinejnyje zadači teorii tonkich plastin, /russisch/, Vestnik Leningr. universiteta, 19, 1958.
5. FIFE P., Nonlinear deflection of thin elastic plates under tension, Comm.Pure Appl. Math., 14,81-112, 1961.
6. BERGER M., On von Kármán's equations and the buckling of a thin elastic plate, Comm. Pure Appl. Math., 20, 687 - 719, 1967.
7. DUBINSKIJ J.A., O rozrešimosti sistemy uravnenij siľnogo izgiba plastinok /russisch/; DAN SSSR, 122, 1,1967.
8. KNIGHTLY G.H., An existence theorem for von Kármán equations, Arch. Rational Mech. Anal. 27, 3, 1967.
9. BERGER M., FIFE P., On von Kármán's equations and the buckling of a thin elastic plate, Comm. Pure Appl. Math., 21, 227 - 241, 1968.
10. KODNÁR R., Použiteľnosť niektorých metód na riešenie nelineárnych úloh teorie stien a plochých škrupín, Stav.čas. SAV, 7, 1968.
11. KODNÁR R., Riešiteľnosť systémov algebraických rovníc nelineárnej teorie stien a plochých škrupín, Stav. čas. SAV, 10, 1968.
12. VOROVIČ I.I., Nekotoryje voprosy ustojčivosti oboloček v boľšom, /russisch/, DAN SSSR, 122, 1, 1958.
13. DJUBEK J., KODNÁR R., Riešenie nelineárnych úloh teórie štíhlych stien vyriačnými metodami, Vydávateľstvo SAV, Bratislava, 1965.
14. KODNÁR R., Probleme nichtlinearer Operatoren bei Untersuchungen der Stabilität dünner Platten und Schalen, Instability of Continuous Systems, Springer 1971.

15. KRASNOSEĽSKIJ M.A., Topologičeskije metody v teorii nelinejnych integraľnych uravnenij, /russisch/, Moskva 1956.
16. SMIRNOV V.I., Kurs vyššej matematiki V, /russisch/, Moskva 1960.
17. SOBOLEV S.L., Nekotoryje primenenija funkcionaľnogo analiza v matematičeskoj fizike, /russisch/, Leningrad 1950.
18. YOSIDA K., Funkcionaľnyj analiz, /russisch/, Moskva 1967.
19. VAJNBERG M.M., Vyriacionnyje metody issledovanija nelinejnych operatorov, /russisch/, Moskva 1956.
20. MICHLIN S.G., Kurs matematičeskij fiziki, /russisch/ Moskva 1968.
21. VOROVIČ I.I., Nekotoryje voprosy ustojčivosti oboloček v boľšom, /russisch/, DAN SSSR, 122, . 1, 1958.

Eingegangen am 27.Mai 1971, in die Redaktion am 5. 1. 1972

Adresse des Autors: Katedra matematickej analyzy PFUK, Bratislava
Matematický pavilón, Mlynská dolina.

(ACTA F. R. N. UNIV. COMEN. - MATHEMATICA XXVIII - 1972)

ACTA FACULTATIS RERUM NATURALIUM UNIVERSITATIS COMENIANAE

MATHEMATICA XXVIII - 1972

CONTRIBUTION TO THE NUMERICAL SOLUTION OF DIFFERENTIAL EQUATIONS BY MEANS OF RUNGE-KUTTA FORMULAS WITH NEWTON-COTES NUMBERS WEIGHTS

ANTON HUŤA, Bratislava

The contents of this paper is the numerical solution of an ordinary differential equation of the 1st order by means of RUNGE - - KUTTA method, by which the parameters are rational numbers and the weights are the NEWTON-COTES numbers. The formulas of the 5th order are derived en detail and also are given the relative simple parameters which are necessary to the construction of formulas of the 6th order. Finally there are given the formulas of the 6th order.

1. Introduction. KUTTA constructed for numerical solution of differential equation

$$y' = f(x,y) \tag{1.1}$$

besides the well-known RUNGE-KUTTA formulas also other ones of the 3rd and 4th order [1] . The formulas of the 3rd order have the form (until the denotation)

$$\begin{aligned} k_0 &= h.f(x_0 \quad ,y_0) \\ k_1 &= h.f(x_0 + \tfrac{1}{2}h, \quad y_0 + \tfrac{1}{2} k_0) \\ k_2 &= h.f(x_0 + h \quad ,y_0 - k_0 + 2k_1) \\ k &= \tfrac{1}{6}(k_0 + 4k_1 + k_2) \end{aligned} \tag{1.2}$$

and the ones of the 4th order

$$\begin{aligned} k_0 &= h.f(x_0 \quad ,y_0) \\ k_1 &= h.f(x_0 + \tfrac{1}{3} h \ ,y_0 + \tfrac{1}{3} k_0) \\ k_2 &= h.f(x_0 + \tfrac{2}{3} h \ ,y_0 - \tfrac{1}{3} k_0 + k_1) \\ k_3 &= h.f(x_0 + \quad h \ ,y_0 + k_0 - k_1 + k_2) \\ k &= \tfrac{1}{8}(k_0 + 3k_1 + 3k_2 + k_3) \end{aligned} \tag{1.3}$$

The weights (i.e.the coefficients k_i for i = 0,1,2,3) in the expressions for k in the formulas (1.2) and (1.3) are the Newton-Cotes numbers.

The object of this contribution is to derive analogical formulas for the 5th and 6th order.

2. The NEWTON-COTES formula for the numerical quadrature has the form [2]

$$\int_a^b f(x)dx = (b-a) \sum_{k=0}^{n} B_k^n f(a+kh') \quad ^{1)} \qquad (2.1)$$

where $h = \frac{b-a}{n}$ is the step and B_k^n is the k-th NEWTON-COTES' number of the n-th order. By choosing a=0, b=1 we have $h' = \frac{1}{n}$. If a_k denote $a+kh' = \frac{k}{n}$ then by choosing $f(x) = x^s$ the formula (2.1) changes into

$$\sum_{k=0}^{n} B_k^n \; a_k^s = \int_0^1 x^s dx = \frac{1}{s+1} \qquad (2.2).$$

As by the division of the interval <a,b> into n parts, the line $y = f(x)$ passes by n points holds the formula (2.2) for s = 0,1,2,...,n.

3. D e n o t a t i o n. During this paper we shall are referring to the paper [3] with the exception that we are changing some of the parameters. (The old denominations will be there in parenthesis.)

m_0, m_q, a_q (φ_q), λ_q (l_q) for q = 1,2,3,4,5,6,7

b_{11} (α), b_{21} (β_0), b_{22} (β_1), b_{31} (γ_0), b_{32} (γ_1), b_{33} (γ_2),

$b_{4,q}(\delta_{q-1})$ for q = 1,2,3,4 $b_{5,q}$ (ε_{q-1}) for q = 1,2,3,4,5

$b_{6,q}$ (ζ_{q-1}) for q = 1,2,3,4,5,6 $b_{7,q}$ (η_{q-1}) for q=1,2,3,4,5,6,7

$c_{1,i}(u_{i-1})$, $c_{2,i}(v_{i-1})$, $c_{3,i}(w_{i-1})$, $c_{4,i}(t_{i-1})$, $c_{5,i}$ for i=2,3,4,5,6,7

$d_{1,i}(s_{i-2})$, $d_{2,i}(r_{i-2})$, $d_{3,i}(q_{i-2})$, $d_{4,i}$ for i = 3,4,5,6,7

$e_{1,i}(T_{i-3})$, $e_{2,i}(R_{i-3})$, $e_{3,i}$ for i = 4,5,6,7 $g_{1,i}$ i = 5,6,7

$d_{1,i}^{(1)}(U_{i-2})$, $d_{2,i}^{(1)}(V_{i-2})$, $d_{1,i}^{(2)}$ for i = 3,4,5,6,7

E_i for i = 3,4,5,6,7 L_i (z_i) for i = 1,2,3,4,5,6.

1/ The denotation h' is given in order to accentuate the difference with h in the following text.

The relations between the parameters

$$a_i = \sum_{j=1}^{i} b_{i,j} \quad \text{for } i = 1,2,3,4,5,6,7 \text{ (the so called RUNGE's conditions)}$$

$$c_{k,i} = \sum_{j=2}^{i} a_{j-1}^{k} b_{i,j} \quad \text{for } i = 2,3,4,5,6,7 \quad k = 1,2,3,4$$

$$d_{k,i} = \sum_{j=3}^{i} c_{k,j-1} b_{i,j} \quad \text{for } i = 3,4,5,6,7 \quad k = 1,2,3$$

$$e_{k,i} = \sum_{j=4}^{i} d_{k,j-1} b_{i,j} \quad \text{for } i = 4,5,6,7 \quad k = 1,2$$

$$g_{1,i} = \sum_{j=5}^{i} e_{1,j-1} b_{i,j}$$

$$d_{k,i}^{(1)} = \sum_{j=3}^{i} a_{j-1}^{\ell} c_{k,j-1} b_{i,j} \quad \text{for } i = 3,4,5,6,7 \; k = 1,2 \; \ell = 1,2$$

$$E_i = \sum_{j=3}^{i} c_{1,j-1}^{2} b_{i,j} \quad \text{for } i = 3,4,5,6,7$$

$$e_{1,i}^{(1)} = \sum_{j=4}^{i} a_{j-1} d_{1,j-1} b_{i,j} \quad \text{for } i = 4,5,6,7$$

$$\mathcal{E}_{1,i}^{(1)} = \sum_{j=4}^{i} d_{1,j-1}^{(1)} b_{i,j} \quad \text{for } i = 4,5,6,7$$

$$L_i = \sum_{j=i+1}^{7} p_j b_{j,i+1} \quad \text{for } i = 1,2,3,4,5,6$$

$$M_i = \sum_{j=i+1}^{6} L_j b_{j,i+1} \quad \text{for } i = 1,2,3,4,5$$

4. Let us suppose that the particular solution of the equation (1.1) is

$$y = F(x) \tag{4.1}$$

If is given the initial condition

$$y_o = F(x_o) \tag{4.2}$$

then the increment of the function $F(x)$ for the increment of the variable x is

$$k = \sum_{i=1}^{\infty} \frac{h^i}{i!} F^{(i)}(x_o) \tag{4.3}$$

because holds

$$y' = F'(x) = f(x,y) \tag{4.4}$$

we can write

$$k = \sum_{i=1}^{\infty} \frac{h^i}{i!} f^{(i-1)}(x_o, y_o) \tag{4.5}$$

For the simplification of the derivative of the function $f(x,y)$ it is advisable to introduce some symbols. If we denote $\frac{\partial^{p+q} f}{\partial x^p \partial y^q} = {}^p f_q$ we can introduce the operators.

$$D^n f = \sum_{j=o}^{n} \binom{n}{j} f^j \cdot {}_{n-j} f_j \tag{4.6}$$

$$D^n f_i = \sum_{j=o}^{n} \binom{n}{j} f^j \cdot {}_{n-j} f_{i+j} \tag{4.7}$$

It is easy to show that for the derivative of the operators hold the relations

$$(D^n f)' = D^{n+1} f + n D^{n-1} f_1 \cdot Df \tag{4.8}$$

$$(D^n f_i)' = D^{n+1} f_i + n D^{n-1} f_{i+1} \cdot Df \tag{4.9}$$

Using the relations (4.6), (4.7), (4.8) and (4.9) we can write

$$m = f + \frac{1}{2} Df.h + \frac{1}{6}(D^2f+f_1Df).h^2 + \frac{1}{24}(D^3f+f_1D^2f+f_1^2Df+3Df\,Df_1)h^3 +$$
$$+ \frac{1}{120}[D^4f+f_1D^3f+f_1^2D^2f+f_1^3Df+4D^2fDf_1 + 6D^2f_1Df+7f_1Df_1Df +$$
$$+ 3f_2(Df)^2]h^4 + \frac{1}{720}[D^5f+f_1D^4f+f_1^2D^3f+f_1^3D^2f+f_1^4Df + 5D^3fDf_1 +$$
$$+ 9f_1D^2fDf_1+12f_1^2DfDf_1+ 10D^3f_1Df+16f_1D^2f_1Df+10D^2f_1D^2f+10f_2D^2fDf +$$
$$+13f_1f_2(Df)^2 + 15Df_2(Df)^2 + 15(Df_1)^2Df]h^5 + \ldots \qquad (4.10)$$

where $$k = m \cdot h \qquad (4.11)$$

In order to find an approximate solution we must apply the values of the expressions .

$$m_o = f(x_o, y_o)$$

$$m_q = f(x_o + a_q h,\ y_o + \lambda_q h) \quad \text{for} \quad q = 1,2,3,\ldots, s \qquad (4.12)$$

where

$$a_q = \sum_{j=1}^{q} b_{qj} \qquad (4.13)$$

and

$$\lambda_q = \sum_{j=1}^{q} b_{qj}\ m_{j-1} \qquad (4.14)$$

If we denote stil $$k_i = m_i h \quad \text{for } i = 0,1,2,\ldots, s \qquad (4.15)$$

the increment will be

$$k = \sum_{i=o}^{s} p_i\ k_i \qquad (4.16)$$

or

$$m = \sum_{i=o}^{s} p_i\ m_i \qquad (4.17)$$

Developping the functions in (4.12) into TAYLOR's series we obtain

$$m_q = f(x_o+a_q h, y_o+ \lambda_q h) = f + \sum_{i=1}^{s} \frac{h^i}{i!} \sum_{j=o}^{i} \binom{i}{j} a_q^{i-j} \lambda_q^{j} \cdot {}_{i-j}f_j \qquad (4.18)$$

by the construction of the expression (4.17) and by the comparison the coefficients of the corresponding terms of (4.10) we obtain the system of non - linear equations for p_i , a_q , b_{qj}.

5. F o r m u l a s of the 5th order. In this case s = 5 and the expressions (4.12) till (4.17) are gradually changing into

$$m_o = f(x_o , y_o)$$

$$m_q = f(x_o + a_q h, y_o + \lambda_q h) \quad \text{for} \quad q = 1,2,3,4,5 \tag{5.1}$$

$$a_q = \sum_{j=1}^{q} b_{qj} \tag{5.2}$$

$$\lambda_q = \sum_{j=1}^{q} b_{qj}\, m_{j-1} \tag{5.3}$$

$$k_i = m_i h \quad \text{for} \quad i = 0,1,2,3,4,5 \tag{5.4}$$

$$k = \sum_{i=o}^{5} p_i \, k_i \tag{5.5}$$

$$m = \sum_{i=o}^{5} p_i \, m_i \tag{5.6}$$

$$m_q = f + \sum_{i=1}^{5} \frac{h^i}{i!} \sum_{j=o}^{i} \binom{i}{j} a_q^{i-j} \lambda_q^j \, f_{\cdot\, i-j\, j} \quad \text{where } q=1,2,3,4,5 \tag{5.7}$$

m_q for $q = 1,2,3,4,5$ obtains the values

$$m_1 = f + \sum_{i=1}^{5} \frac{1}{i!} a_1^i D^i f \,.\, h^i \tag{5.8}$$

$$m_2 = f + a_2 Df.h + \tfrac{1}{2}(a_2^2 D^2 f + 2c_{12} f_1 Df).h^2 + \tfrac{1}{6}(a_2^3 D^3 f + 3c_{22} f_1 D^2 f + 6a_2 c_{12} Df Df_1).h^3 + \tfrac{1}{24}[a_2^4 D^4 f + 4c_{32} f_1 D^3 f + 12 a_2^2 c_{12} Df D^2 f_1 + 12 c_{12}^2 f_2 (Df)^2]\, h^4 + \ldots \tag{5.9}$$

$$m_3 = f + a_3 Df.h + \tfrac{1}{2}(a_3^2 D^2 f + 2c_{13} f_1 Df).h^2 + \tfrac{1}{6}(a_3^3 D^3 f + 3c_{23} f_1 D^2 f +$$
$$+ 6d_{13} f_1^2 Df + 6a_3 c_{13} DfDf_1).h^3 + \tfrac{1}{24}[a_3^4 D^4 f + 4c_{33} f_1 D^3 f + 12d_{23} f_1^2 D^2 f +$$
$$+ 24 d_{13}^{(1)} f_1 DfDf_1 + 12a_3 c_{23} D^2 f Df_1 + 24 a_3 d_{13} f_1 DfDf_1 + 12c_{13}^2 f_2 (Df)^2 +$$
$$+ 12a_3^2 c_{13} DfD^2 f_1] h^4 + \ldots \qquad (5.10)$$

$$m_4 = f + a_4 Df.h + \tfrac{1}{2}(a_4^2 D^2 f + 2c_{14} f_1 Df)\, h^2 + \tfrac{1}{6}(a_4^3 D^3 f + 3c_{24} f_1 D^2 f +$$
$$+ 6d_{14} f_1^2 Df + 6a_4 c_{14} DfDf_1)\, h^3 + \tfrac{1}{24}[a_4^4 D^4 f + 4c_{34} f_1 D^3 f + 12d_{24} f_1^2 D^2 f +$$
$$+ 24 e_{14} f_1^3 Df + 24\, d_{14}^{(1)} + a_4 d_{14}) f_1 DfDf_1 + 12a_4 c_{24} D^2 f Df_1 +$$
$$+ 12a_4^2 c_{14} DfD^2 f_1 + 12_{14}^2 f_2 (Df)^2] h^4 + \ldots \qquad (5.11)$$

$$m_5 = f + a_5 Df.h + \tfrac{1}{2}(a_5^2 D^2 f + 2c_{15} f_1 Df).h^2 + \tfrac{1}{6}(a_5^3 D^3 f + 3c_{25} f_1 D^2 f + 6d_{15} f_1^2 Df +$$
$$+ 6a_5 c_{15} DfDf_1).h^3 + \tfrac{1}{24}[a_5^4 D^4 f + 4c_{35} f_1 D^3 f + 12d_{25} f_1^2 D^2 f + 24 e_{15} f_1^3 Df +$$
$$+ 24 (d_{15}^{(1)} + a_5 d_{15}) f_1 DfDf_1 + 12a_5 c_{25} D^2 f Df_1 + 12a_5^2 c_{15} DfD^2 f_1 +$$
$$+ 12c_{15}^2 f_2 (Df)^2] h^4 + \ldots \qquad (5.12)$$

In putting m_o and m_q for $q = 1,2,3,4,5$ and by the comparison of the coefficients of the corresponding terms (4.10) we obtain the system

$$[f] \sum_{i=o}^{5} p_i = 1 \qquad (5.13) \qquad [Df] \sum_{i=1}^{5} p_i a_i = \tfrac{1}{2} \qquad (5.14)$$

$$[D^2 f] \sum_{i=1}^{5} p_i a_i^2 = \tfrac{1}{3} \qquad (5.15) \qquad [D^3 f] \sum_{i=1}^{5} p_i a_i^3 = \tfrac{1}{4} \qquad (5.16)$$

$$[D^4 f] \sum_{i=1}^{5} p_i a_i^4 = \tfrac{1}{5} \qquad (5.17) \qquad [f_1 Df] \sum_{i=2}^{5} p_i c_{1,i} = \tfrac{1}{6} \qquad (5.18)$$

$$[f_1 D^2 f] \sum_{i=2}^{5} p_i c_{2,i} = \tfrac{1}{12} \qquad (5.19) \qquad [f_1 D^3 f] \sum_{i=2}^{5} p_i c_{3,i} = \tfrac{1}{20} \qquad (5.20)$$

$$[DfDf_1] \sum_{i=2}^{5} p_i a_i c_{1,i} = \frac{1}{8} \quad (5.21) \qquad [D^2fDf_1] \sum_{i=2}^{5} p_i a_i c_{2,i} = \frac{1}{15} \quad (5.22)$$

$$[DfD^2f_1] \sum_{i=2}^{5} p_i a_i^2 c_{1,i} = \frac{1}{10} \quad (5.23) \qquad [f_2(Df)^2] \sum_{i=2}^{5} p_i c_{1,i}^2 = \frac{1}{20} \quad (5.24)$$

$$[f_1^2Df] \sum_{i=3}^{5} p_i d_{1,i} = \frac{1}{24} \quad (5.25) \qquad [f_1^2D^2f] \sum_{i=3}^{5} p_i d_{2,i} = \frac{1}{60} \quad (5.26)$$

$$[f_1DfDf_1] \sum_{i=3}^{5} p_i d_{1,i}^{(1)} + \sum_{i=3}^{5} p_i a_i d_{1,i} = \frac{7}{120} \quad (5.27)$$

$$[f_1^3Df] \sum_{i=4}^{5} p_i e_{1,i} = \frac{1}{120} \quad (5.28)$$

where in the brackets are always indicated the terms, the coefficients of which have been compared.

It is a question of the system of 16 equations with 21 unknows i.e. p_i for $i = 0,1,2,3,4,5$; a_i for $i = 1,2,3,4,5$, b_{ij} for $i = 2,3,4,5$ and $j = 2,3,4,5$ $i \neq j$. We can choose of this 5 parameter. If we take $p_1 = 0$, $a_2 = \frac{1}{4}$, $a_3 = \frac{2}{4}$, $a_4 = \frac{3}{4}$, $a_5 = 1$ then the equations (5.14), (5.15), (5.16) and (5.17) are fulfilled, if provided that on the basis of (22) the corresponding weigths have the following values $p_2 = \frac{32}{90}$, $p_3 = \frac{12}{90}$, $p_4 = \frac{32}{90}$, $p_5 = \frac{7}{90}$. As can be seen they are NEWTON - COTES numbers. From the equation (5.13) follows $p_o = \frac{7}{90}$.

If we put into the equations (5.14), (5.15), (5.16) and (5.17) $p_1 = 0$ we obtain the equations (5.14.1), (5.15.1), (5.16.1) and (5.17.1) of the form $\sum_{i=2}^{5} p_i a_i^k = \frac{1}{k+1}$ for $k = 1,2,3,4$.

If we compare the equations (5.15.1), (5.16.1) and (5.17.1) with the equations (5.18), (5.21) and (5.23) we obtain the relations $c_{1,i} = \frac{1}{2} a_i^2$ for $i = 2,3,4,5$ (5.29) ÷ (5.32) and at the same time

is fulfilled the equation (5.24). By comparing the equations (5.16.1) and (5.17.1) with the equations (5.19) and (5.22) we obtain

$$c_{2,i} = \frac{1}{3} a_i^3 \quad \text{for} \quad i = 2,3,4,5 \qquad (5.33 \div 5.36)$$

By the introduction of the substitutions

$$L_i = \sum_{j=i+1}^{5} p_j b_{j,i+1} \quad \text{for} \quad i = 1,2,3,4 \qquad (5.37)$$

into the equations (5.18), (5.19), (5.20), (5.25), (5.26) and (5.28) we obtain the equations

$$\sum_{i=1}^{4} a_i L_i = \frac{1}{6} \quad (5.18.1) \qquad \sum_{i=1}^{4} a_i^2 L_i = \frac{1}{12} \qquad (5.19.1)$$

$$\sum_{i=1}^{4} a_i^3 L_i = \frac{1}{20} \quad (5.20.1) \qquad \sum_{i=2}^{4} c_{1,i} L_i = \frac{1}{24} \qquad (5.25.1)$$

$$\sum_{i=2}^{4} c_{2,i} L_i = \frac{1}{60} \quad (5.26.1) \qquad \sum_{i=3}^{4} d_{1,i} L_i = \frac{1}{120} \qquad (5.28.1)$$

From the relations (5.19.1) - 2. (5.25.1) and at the same time from (5.20.1) - 2. (5.26.1) follows $L_1 = 0$.

From this supposition the system of equations (5.18.1), (5.19.1) and (5.20.1) has a solution $L_2 = \frac{12}{45}$, $L_3 = \frac{3}{45}$, $L_4 = \frac{4}{45}$.

By the introduction of the substitutions

$$M_i = \sum_{j=i+1}^{3} L_j b_{j,i+1} \quad \text{for} \quad i = 1,2,3 \qquad (5.38)$$

into the equations (5.25.1), (5.26.1) and (5.28.1)

we obtain

$$\sum_{i=1}^{3} a_i M_i = \frac{1}{24} \quad (5.25.2) \qquad \sum_{i=1}^{3} a_i^2 M_i = \frac{1}{60} \quad (5.26.2)$$

$$\sum_{i=2}^{3} c_{1,i} M_i = \frac{1}{120} \quad (5.28.2)$$

of which for $a_1 \neq 0$ follows $M_1 = 0$, $M_2 = \frac{1}{15}$, $M_3 = \frac{1}{20}$
From p_5 and L_4 follows $b_{55} = \frac{8}{7}$, from L_4 and M_3 we have
$b_{44} = \frac{9}{16}$; from L_3 we obtain $b_{54} = -\frac{12}{7}$.

Finally from the systems $c_{1,i}$, $c_{2,i}$ for $i = 2,3,4,5$ we obtain gradually $a_1 = \frac{1}{6}$, $b_{22} = \frac{3}{16}$, $b_{32} = -\frac{3}{4}$, $b_{33} = 1$ $b_{42} = 0$, $b_{43} = 0$, $b_{52} = \frac{3}{7}$, $b_{53} = \frac{12}{7}$.

From the RUNGE's conditions follows $b_{11} = \frac{1}{6}$, $b_{21} = \frac{1}{16}$, $b_{31} = \frac{1}{4}$, $b_{41} = \frac{3}{16}$, $b_{51} = -\frac{4}{7}$.

The formulas of the 5th order have the from

$$k_o = f\ (x_o\ ,\ y_o)\ .\ h$$

$$k_1 = f\ (x_o + \tfrac{1}{6} h,\ y_o + \tfrac{1}{6} k_o)\ .\ h$$

$$k_2 = f\ (x_o + \tfrac{1}{4} h,\ y_o + \frac{k_o+3k_1}{16})\ .\ h$$

$$k_3 = f\ (x_o + \tfrac{1}{2} h,\ y_o + \frac{k_o-3k_1+4k_2}{4})\ .\ h$$

$$k_4 = f\ (x_o + \tfrac{3}{4} h,\ y_o + \frac{3k_o+9k_3}{16})\ .\ h$$

$$k_5 = f\ (x_o + h,\ y_o + \frac{-4k_o+3k_1+12k_2-12k_3+8k_4}{7})\ .\ h$$

$$k = \frac{1}{90}\ (7k_o + 32k_2+12k_3+32k_4+7k_5)$$

6. F o r m u l a s of the 6th order. There is s = 7. In this section we shall direct our attention to the paper [3] and we shall consider only the results which are different from the results of the mentioned paper.

$$\sum_{i=o}^{7} p_i = 1 \qquad (6.1)$$

$$\sum_{i=1}^{7} p_i a_i^k = \frac{1}{k+1} \quad \text{for} \quad k = 1,2,3,4,5 \qquad (6.2) \div (6,6)$$

$$\sum_{i=2}^{7} p_i c_{k,i} = \frac{1}{(k+1)(k+2)} \quad \text{for} \quad k = 1,2,3,4 \qquad (6.7) \div (6.10)$$

$$\sum_{i=2}^{7} p_i a_i c_{k,i} = \frac{1}{(k+1)(k+3)} \quad \text{for} \quad k = 1,2,3 \qquad (6.11) \div (6.13)$$

$$\sum_{i=2}^{7} p_i a_i^2 c_{k,i} = \frac{1}{(k+1)(k+4)} \quad \text{for} \quad k = 1,2 \qquad (6.14) \div (6,15)$$

$$\sum_{i=2}^{7} p_i a_i^3 c_{1,i} = \frac{1}{12} \qquad (6.16)$$

$$\sum_{i=2}^{7} p_i a_i^k c_{1,i} = \frac{1}{4(k+5)} \quad \text{for} \quad k = 0,1 \qquad (6.17) \div (6.18)$$

$$\sum_{i=2}^{7} p_i c_{1,i} c_{2,i} = \frac{1}{36} \qquad (6.19)$$

$$\sum_{i=3}^{7} p_i d_{k,i} = \frac{1}{(k+1)(k+2)(k+3)} \quad \text{for} \quad k = 1,2,3 \qquad (6.20) \div (6.22)$$

$$\sum_{i=3}^{7} p_i d_{1,i}^{(1)} + \sum_{i=3}^{7} p_i a_i d_{1,i} = \frac{7}{120} \qquad (6.23)$$

$$\sum_{i=3}^{7} p_i d_{2,i}^{(1)} + \sum_{i=3}^{7} p_i a_i d_{2,i} = \frac{1}{40} \qquad (6.24)$$

$$\sum_{i=3}^{7} p_i d_{1,i}^{(2)} + \sum_{i=3}^{7} p_i a_i^2 d_{1,i} = \frac{2}{45} \qquad (6.25)$$

$$\sum_{i=3}^{7} p_i a_i d_{1,i}^{(1)} = \frac{1}{48} \qquad (6.26)$$

$$2 \sum_{i=3}^{7} p_i c_{1,i} d_{1,i} + \sum_{i=3}^{7} p_i E_i = \frac{13}{360} \qquad (6.27)$$

$$\sum_{i=4}^{7} p_i e_{k,i} = \frac{1}{(k+1)(k+2)(k+3)(k+4)} \text{ for } k = 1,2 \qquad (6.28) \text{ \& } (6.29)$$

$$\sum_{i=4}^{7} p_i \varepsilon_{1,i}^{(1)} + \sum_{i=4}^{7} p_i e_{1,i}^{(1)} + \sum_{i=4}^{7} p_i a_i e_{1,i} = \frac{1}{60} \qquad (6.30)$$

$$\sum_{i=5}^{7} p_i g_{1,i} = \frac{1}{720} \qquad (6.31)$$

By the introduction of the substitutions

$$L_i = \sum_{j=i+1}^{7} p_j b_{j,i+1} \quad \text{for} \quad i = 1,2,3,4,5,6 \qquad (6.32)$$

into the equations (6.7), (6.8), (6.9), (6.10), (6.20), (6.21), (6.22), (6.28), (6.29) and (6.31) we obtain the system

$$\sum_{i=1}^{6} a_i^k L_i = \frac{1}{(k+1)(k+2)} \quad \text{for} \quad k = 1,2,3,4 \qquad (6.7.1) \text{ to } (6.10.1)$$

$$\sum_{i=2}^{6} c_{k,i} L_i = \frac{1}{(k+1)(k+2)(k+3)} \quad \text{for} \quad k = 1,2,3 \qquad (6.20.1) \text{ to } (6.22.1)$$

$$\sum_{i=3}^{6} d_{k,i} L_i = \frac{1}{(k+1)(k+2)(k+3)(k+4)} \quad \text{for } k = 1,2 \qquad (6.28.1) \text{ and } (6.29.1)$$

$$\sum_{i=4}^{6} e_{1,i} L_i = \frac{1}{720} \qquad (6.31.1)$$

and at last by the introduction of the substitutions

$$M_i = \sum_{j=i+1}^{6} L_j b_{j,i+1} \quad \text{for} \quad i = 1,2,3,4,5 \qquad (6.33)$$

into the equations (6.20.1), (6.21.1), (6.22.1), (6.28.1), (6.29.1) and (6.31.1) we have

$$\sum_{i=1}^{5} a_i^k M_i = \frac{1}{(k+1)(k+2)(k+3)} \quad \text{for } k = 1,2,3 \qquad (6.20.2) \text{ to } (6.22.2)$$

$$\sum_{i=2}^{5} c_{k,i} M_i = \frac{1}{(k+1)(k+2)(k+3)(k+4)} \quad \text{for } k=1,2 \qquad (6.28.2) \text{ and } (6.29.2)$$

$$\sum_{i=3}^{5} d_{1,i} M_i = \frac{1}{720} \qquad (6.31.2)$$

One solution of the above mentioned systems is given

$a_1 = \frac{1}{10}$, $a_2 = \frac{1}{5}$, $a_3 = \frac{1}{5}$, $a_4 = \frac{2}{5}$, $a_5 = \frac{3}{5}$, $a_6 = \frac{4}{5}$, $a_7 = 1$

$p_0 = p_7 = \frac{19}{288}$, $p_1 = p_2 = 0$, $p_3 = p_6 = \frac{75}{288}$, $p_4 = p_5 = \frac{50}{288}$

$L_1 = 0$, $L_2 = 0$, $L_3 = \frac{5}{24}$, $L_4 = \frac{5}{48}$, $L_5 = \frac{5}{72}$, $L_6 = \frac{5}{96}$, $M_1 = 0$, $M_2 = 0$

$M_3 = \frac{5}{48}$, $M_4 = 0$, $M_5 = \frac{5}{144}$, $c_{12} = \frac{1}{25}$, $c_{13} = \frac{1}{50}$, $c_{14} = \frac{4}{50}$, $c_{15} = \frac{9}{50}$,

$c_{16} = \frac{16}{50}$, $c_{17} = \frac{25}{50}$, $c_{22} = \frac{1}{250}$, $c_{23} = \frac{1}{375}$, $c_{24} = \frac{8}{375}$, $c_{25} = \frac{27}{375}$, $c_{26} = \frac{64}{375}$,

$c_{27} = \frac{125}{375}$, $d_{13} = \frac{1}{750}$, $d_{14} = \frac{8}{750}$, $d_{15} = \frac{27}{750}$, $d_{16} = \frac{64}{750}$, $d_{17} = \frac{125}{750}$,

$e_{14} = \frac{4}{3750}$, $e_{15} = \frac{18}{3750}$, $e_{16} = \frac{68}{3750}$, $e_{17} = \frac{560}{19.750}$, $g_{15} = \frac{8}{18750}$

$g_{16} = \frac{148}{56250}$, $g_{17} = \frac{680}{19.3750}$

$b_{11} = \frac{1}{10}$, $b_{21} = -\frac{1}{5}$, $b_{31} = \frac{1}{30}$, $b_{41} = \frac{4}{15}$, $b_{51} = \frac{3}{10}$, $b_{61} = -\frac{8}{15}$, $b_{71} = \frac{169}{114}$,

$b_{22} = \frac{2}{5}$, $b_{32} = \frac{2}{15}$, $b_{42} = -\frac{8}{15}$, $b_{52} = -\frac{6}{15}$, $b_{62} = \frac{16}{15}$, $b_{72} = -\frac{130}{57}$, $b_{33} = \frac{1}{30}$

$b_{43} = -\frac{4}{30}$, $b_{53} = \frac{-3}{30}$, $b_{63} = \frac{8}{30}$, $b_{73} = -\frac{65}{114}$, $b_{44} = \frac{4}{5}$, $b_{54} = \frac{2}{5}$, $b_{64} = -\frac{2}{15}$

$b_{74} = \frac{10}{19}$, $b_{55} = \frac{2}{5}$, $b_{65} = -\frac{8}{15}$, $b_{75} = \frac{50}{19}$, $b_{66} = \frac{2}{3}$, $b_{76} = -\frac{30}{19}$, $b_{77} = \frac{15}{19}$

So the formulas of the 6th order will have the following form

$k_0 = f(x_0 \quad , \quad y_0) \cdot h$

$k_1 = f(x_0 + \frac{1}{10} h \ , \ y_0 + \frac{k_0}{10}) \cdot h$

$k_2 = f(x_0 + \frac{1}{5} h \ , \ y_0 + \frac{-k_0 + 2k_1}{5}) \cdot h$

$k_3 = f(x_0 + \frac{1}{5} h \ , \ y_0 + \frac{k_0 + 4k_1 + k_2}{30}) \cdot h$

$$k_4 = f(x_0 + \frac{2}{5}h\ ,\ y_0 + \frac{4k_0 - 8k_1 - 2k_2 + 12k_3}{15}).h$$

$$k_5 = f(x_0 + \frac{3}{5}h\ ,\ y_0 + \frac{3k_0 - 4k_1 - k_2 + 4k_3 + 4k_4}{10}).h$$

$$k_6 = f(x_0 + \frac{4}{5}h\ ,\ y_0 + \frac{-8k_0+16k_1+4k_2-2k_3-8k_4+10k_5}{15}).h$$

$$k_7 = f(x_0 + h\ ,\ y_0 + \frac{169k_0-260k_1-65k_2+60k_3+300k_4-180_5+90k_6}{114}).h$$

$$k = \frac{19k_0+75k_3+50k_4+50k_5+75k_6+19k_7}{288}$$

References

[1] KOPAL.Z.: Numerical Analysis 1955. Pg.204 and 209.

[2] KRYLOV.V.J.: Približennoe vyčíslenie integralov. Moskva 1959. Pg.79.

[3] HUŤA .A.: Une amélioration de la méthode de RUNGE-KUTTA-NYSTRÖM pour la résolution numérique des équations différentielles du premier ordre. Acta facultatis Rerum naturalium Universitatis Comenianae. Bratislava 1956.

Received October 1, 1971, so the publishing January 5, 1972

Author´s address: Katedra numerickej matematiky a matematickej štatistiky PFUK, Bratislava,
Mlynská dolina, Matematický pavilón

(ACTA F. R. N. UNIV. COMEN. - MATHEMATICA XXVIII - 1972)
ACTA FACULTATIS RERUM NATURALIUM UNIVERSITATIS COMENIANAE
MATHEMATICA XXVIII - 1972

AN ERROR ESTIMATION OF A HUŤA FORMULA OF THE RUNGE-KUTTA TYPE OF THE SIXTH ORDER

ANNA BERECOVÁ, Bratislava

In the paper is derived an error estimation of the approximate solution of first-order diferential equation by the Runge-Kutta type method of the sixth order.

Introduction

In this paper is derived an error estimation of the approximate solution of first-order differential equation by the Runge-Kutta type method of the sixth order /see HUŤA [2] /.

My approach is analogical to that used by VEJVODA O. [3] for the error estimation of Runge-Kutta formula of the fourth order and by JUKL V. [4] for Nyström formula of the 5-th order.

The paper has three sections. In the first part the brief description of the Huťa formula is presented. In the second part the error estimation is given when one step is done. I remark that the estimation is essentially better by Huťa´s formula of the sixth order as in the case of the 4-th and 5-th order formulas. In the third part the formula for the error estimation, when s steps are done, is derived.

The brief description of the HUŤA formula

Given the first-order differential equation

$$y' = f(x,y) \qquad /1/$$

with the initial condition

$$y(x_0) = y_0 \qquad /2/$$

it is assumed that $f(x,y) \in C^6(B)$, where $B = \langle x_0, X\rangle \times \langle y_0-b, y_0+b\rangle$, $X > x_0$. It is furter assumed that the LIPSCHITZ condition with constant K is fulfilled in the region B

$$|f(x,y_1) - f(x,y_2)| \leq K|y_1 - y_2| \qquad /3/$$

Put $A = \max_{(x,y) \in B} |f(x,y)|$ and let $|X - x_0| A \leq b$.

These conditions guarantee the existence of a continuous function $y = y(x)$ such that for $x \in \langle x_0, X\rangle$ $y(x)$ satisfies /1/, /2/; $y(x)$ is uniquely determined by these conditions. Moreover $y(x)$ is differentiable up to order 7.

Let $\bar{y}(x)$ be the approximate solution of /1/, /2/ such that its Taylor expansion about the point (x_0,y_0) agrees with the exact solution up to terms of order 6. We suppose that $\bar{y}(x)$ can be written in the form

$$\bar{y}(x) = y_0 + K \qquad /5/$$

where

$$K = \Delta\bar{y} = h \sum_{i=0}^{7} p_i m_i \qquad /6/$$

$$h = \Delta x = x - x_0 \ , \quad 0 < h < X - x_0$$

p_i are constants for $i = 0,1,2,\ldots,7$ and m_i are the values of $f(x,y)$ defined as follows

$$m_0 = f(x_0,y_0)$$

$$m_i = f(x_0 + a_i h \ , \ y_0 + \lambda_i h), \quad i = 1,2,\ldots,7 \qquad /7/$$

Moreover the relations $\lambda_i = \sum_{j=1}^{i} b_{i,j}\, m_{j-1}$ /8/

and $a_i = \sum_{j=1}^{i} b_{ij}$, $i = 1,2,\ldots,7$ /9/

/so called RUNGE-conditions/ hold.

Here is one of the formulas of order 6-th derived by HUTA A. [2] .

$$K = \frac{1}{288}\,(19K_0 + 75K_3 + 50K_4 + 50K_5 + 75K_6 + 19K_7) \qquad /10/$$

where

$$K_0 = f(x_0, y_0)\,h$$

$$K_1 = f\left(x_0 + \frac{1}{10}\,h,\; y_0 + \frac{1}{10}\,K_0\right) h$$

$$K_2 = f\left(x_0 + \frac{1}{5}\,h,\; y_0 + \frac{-K_0 + 2K_1}{5}\right)h$$

$$K_3 = f\left(x_0 + \frac{1}{5}\,h,\; y_0 + \frac{K_0 + 4K_1 + K_2}{30}\right) h$$

$$K_4 = f\left(x_0 + \frac{2}{5}\,h,\; y_0 + \frac{4K_0 - 8K_1 - 2K_2 + 12K_3}{15}\right) h$$

$$K_5 = f\left(x_0 + \frac{3}{5}\,h,\; y_0 + \frac{3K_0 - 4K_1 - K_2 + 4K_3 + 4K_4}{10}\right) h$$

$$K_6 = f\left(x_0 + \frac{4}{5}\,h,\; y_0 + \frac{-8K_0 + 16K_1 + 4K_2 - 2K_3 - 8K_4 + 10K_5}{15}\right) h$$

$$K_7 = f\left(x_0 + h,\; y_0 + \frac{169K_0 - 260K_1 - 65K_2 + 60K_3 + 300K_4 - 180K_5 + 90K_6}{114}\right) h$$

These formulas are interesting, since Newton-Cotes coeficients appear in the expression for K. Hence coeficients a_i, b_{ij} and p_i are as follows:

$$a_1 = \frac{1}{10},\; a_2 = \frac{1}{5},\; a_3 = \frac{1}{5},\; a_4 = \frac{2}{5},\; a_5 = \frac{3}{5},\; a_6 = \frac{4}{5},\; a_7 = 1$$

$$b_{11} = \frac{1}{10}, b_{21} = \frac{-1}{5}, b_{31} = \frac{1}{30}, b_{41} = \frac{4}{15}, b_{51} = \frac{3}{10}, b_{61} = \frac{-8}{15}, b_{71} = \frac{169}{114}$$

$$b_{22} = \frac{2}{5}, b_{32} = \frac{2}{15}, b_{42} = \frac{-8}{15}, b_{52} = \frac{-2}{5}, b_{62} = \frac{16}{15}, b_{72} = \frac{-130}{57}$$

$$b_{33} = \frac{1}{30}, b_{43} = \frac{-2}{15}, b_{53} = \frac{-1}{10}, b_{63} = \frac{4}{15}, b_{73} = \frac{-65}{114}$$

$$b_{44} = \frac{4}{5}, b_{54} = \frac{2}{5}, b_{64} = \frac{-2}{15}, b_{74} = \frac{10}{19}$$

$$b_{55} = \frac{2}{5}, b_{65} = \frac{-8}{15}, b_{75} = \frac{50}{19}$$

$$b_{66} = \frac{2}{3}, b_{76} = \frac{-30}{19}$$

$$b_{77} = \frac{15}{19} \qquad /11/$$

$$p_0 = \frac{19}{288},\; p_1 = p_2 = 0,\; p_3 = \frac{75}{288}, p_4 = p_5 = \frac{50}{288}, p_6 = \frac{75}{288}, p_7 = \frac{19}{288}$$

The error estimation at one step

We define the error of the integration method on the interval $\langle x_0, X\rangle$, $x < X$ committed at each step as follows:

$$\gamma(x_0, x) = \max_{\xi \in \langle x_0, x\rangle} |\bar{y}(\xi) - y(\xi)| \qquad /12/$$

To obtain the estimation of γ we expand both the exact and the approximate solutions to Taylor series up to terms of order 7-th. These expansions are identical up to terms of order 6-th.
The next notation follows /see [1] /

$$f(x,y) = f; \quad \frac{\partial^{p+q} f}{\partial x^p \partial y^q} = {}_pf_q$$

and f_q and ${}_pf$ mean that $p = 0$ or $q = 0$, respectively.
For the functions differentiable up to order 6-th, we define the operations

$$D^n f = \sum_{j=0}^{n} \binom{n}{j} {}_{n-j}f_j \cdot f^j \quad 1 \leqq n \leqq 6 \qquad /13/$$

$$D^n f_k = \sum_{j=0}^{n} \binom{n}{j} {}_{n-j}f_{j+k} \cdot f^j \qquad /14/$$

Differentiating with respect to x yields:

$$(D^n f)' = D^{n+1} f + n D^{n-1} f_1 \, Df \qquad /15/$$

$$(D^n f_k)' = D^{n+1} f_k + n D^{n-1} f_{k+1} \, Df \qquad /16/$$

We now proceed to expand the exact solution $y(x)$ about the point (x_0, y_0). First we have

$$y(x) = y(x_0) + \sum_{i=1}^{7} \frac{h^i}{i!} y^{(i)}(x_0) + O(h^8)$$

Using /1/, /2/, we obtain

$$y(x) = y_o + \sum_{i=1}^{7} \frac{h^i}{i!} \left[f^{(i-1)}\right]_{x=x_o} + O(h^8) \qquad /17/$$

Finally, usin /13/,/14/,/15/,/16/ we obtain

$$y' = f$$

$$f' = {}_1f + f_1.f = Df$$

$$f'' = (Df)' = D^2f + f_1.Df$$

$$f''' = D^3f + 3\,Df_1\,Df + f_1\,D^2f + f_1^2\,Df$$

$$f^{IV} = D^4f + f_1\,D^3f + f_1^2\,D^2f + f_1^3\,Df + 4\,D^2f\,Df_1 + 6D^2f_1\,Df + 7\,f_1\,Df_1\,Df + 3\,f_2\,(Df)^2$$

$$f^{V} = D^5f + f_1D^4f + f_1^2D^3f + f_1^3D^2f + f_1^4Df + 5\,D^3f\,Df_1 + 9\,f_1D^2f\,Df_1 + 12\,f_1^2Df\,Df_1 + 10\,D^3f_1Df + 16\,f_1D^2f_1Df_1 + 10\,D^2f_1D^2f + 10\,f_2\,D^2f\,Df + 13\,f_1f_2(Df)^2 + 15\,Df_2(Df)^2 + 15\,(Df_1)^2\,Df$$

$$f^{VI} = D^6f + f_1D^5f + f_1^2D^4f + f_1^3D^3f + f_1^4D^2f + f_1^5Df + 6\,D^4f\,Df_1 + 11\,f_1\,D^3f\,Df_1 + 15\,f_1^2\,D^2f\,Df_1 + 18\,f_1^3\,Df\,Df_1 + 15\,Df\,D^4f_1 + 30\,f_1\,Df\,D^3f_1 + 31\,f_1^2\,Df\,D^2f_1 + 15\,D^3f\,D^2f_1 + 25\,f_1D^2f\,D^2f_1 + 20\,D^2f\,D^3f_1 + 81\,Df\,Df_1D^2f_1 + 15f_2\,DfD^3f + 45f_1\,f_2Df\,D^2f + 24\,D^2f\,(Df_1)^2 + 57\,f_1\,Df(Df_1)^2 + 45(Df)^2\,D^2f_2 + 75\,f_1(Df)^2\,Df_2 + 60\,Df\,D^2f\,Df_2 + 63\,f_2(Df)^2\,Df_1 + 10\,f_2(D^2f)^2 + 38\,f_1^2\,f_2\,(Df)^2 + 15\,f_3(Df)^3$$

After the substitution into /17/, we obtain the expansion for the exact solution y (x) in the form

$$y\,(x) = y_o + \sum_{i=1}^{6} \frac{h^i}{i!} \left[f^{(i-1)}\right]_{x=x_o} + \frac{h^7}{5040} \left[D^6f + f_1\,D^5f + f_1^2\,D^4f + f_1^3D^3f + f_1^4\,D^2f + f_1^5Df + 6\,D^4f\,Df_1 + 11\,f_1D^3f\,Df_1 + \right.$$

$$+ 15 f_1^2 D^2 f\, Df_1 + 18 f_1^3\, Df\, Df_1 + 15\, Df\, D^4 f_1 + 30 f_1\, Df\, D^3 f_1 +$$
$$+ 31 f_1^2\, Df\, D^2 f_1 + 15\, D^3 f\, D^2 f_1 + 25 f_1\, D^2 f\, D^2 f_1 + 20\, D^2 f\, D^3 f_1 +$$
$$+ 81\, Df\, Df_1\, D^2 f_1 + 15 f_2 Df\, D^3 f + 45 f_1 f_2\, Df\, D^2 f + 24\, D^2 f\, (Df_1)^2 +$$
$$+ 57 f_1 Df (Df_1)^2 + 45 (Df)^2 D^2 f_2 + 75 f_1 (Df)^2 Df_2 + 60\, Df\, D^2 f\, Df_2 +$$
$$+ 63 f_2 (Df)^2\, Df_1 + 10 f_2 (D^2 f)^2 + 38 f_1^2 f_2\, (Df)^2 + 15 f_3\, (Df)^3 +$$
$$+ 0\,(h^8) \qquad /18/$$

We also expand the approximate solution given by formulas /5/, /6/, /7/. We have

$$m_i = \sum_{k=0}^{6} \frac{h^k}{k!} \sum_{j=0}^{k} \binom{k}{j} a_i^{k-j} \lambda_i^j \cdot {}_{k-j}f_j \qquad /19/$$

for $i = 1,2,\ldots,7$.

We are interested only in the terms with h^6 /or h^7 when considering K_i instead of m_i/, since the preceding terms vanish after the substitution into /12/. In particular, from /19/, we have

$$m_1 = f + \ldots + \frac{1}{720}\left[a_1^6\, D^6 f\right] h^6 + \ldots$$

$$m_2 = f + \ldots + \frac{1}{720}\left[a_2^6\, D^6 f + 6\, c_{52}\, f_1\, D^5 f + 30\, a_2\, c_{42} D^4 f\, Df_1 + \right.$$
$$+ 30 a_2^4\, c_{12}\, D f\, D^4 f_1 + 60 a_2^2\, c_{32} D^3 f\, D^2 f_1 + 60 a_2^3\, c_{22} D^2 f\, D^3 f_1 +$$
$$+ 120\, c_{12} c_{32} f_2 Df D^3 f + 180 a_2^2 c_{12}^2 (Df)^2\, D_2 f_2 + 360\, a_2 c_{12} c_{22} Df D^2 f Df_2 +$$
$$\left. + 90\, c_{22}^2\, f_2 (D^2 f)^2 + 120\, c_{12}^3\, f_3 (Df)^3\right] h^6 + \ldots$$

$$m_3 = f + \ldots + \frac{1}{720}\left[a_3^6\, D^6 f + 6\, c_{53} f_1 D^5 f + 30\, d_{43} f_1^2\, D^4 f + \right.$$
$$+ 30\, a_3 c_{43} D^4 f\, Df_1 + 120\, (A_{133} + a_3 d_{33})\, f_1\, D^3 f\, Df_1 +$$
$$\left. + 30 a_3^4 c_{13} Df\, D^4 f_1 + 120\, (A_{313} + a_3^3 d_{13}) f_1 Df D^3 f_1 + 60 a_3^2 c_{33} D^3 f D^2 f_1 + \right.$$

$$+ 180\,(A_{223} + a_3^2 d_{23}) f_1 D^2 f D^2 f_1 + 60\, a_3^3 c_{23} D^2 f D^3 f_1 +$$

$$+ 360 (a_3^2 A_{113} + a_3 A_{213}) Df Df_1 D^2 f_1 + 120\, c_{13} c_{33} f_2\, Df\, D^3 f +$$

$$+ 360\,(c_{13} d_{23} + c_{23} d_{13}) f_1 f_2\, Df D^2 f +$$

$$+ 360\, a_3 A_{123} D^2 f\, (Df_1)^2 + 180\, a_3^2 c_{13}^2 (Df)^2 D^2 f_2 + 360 (2 a_3 c_{13} d_{13} +$$

$$+ E_{13}) f_1 (Df)^2\, Df_2 + 360\, a_3 c_{13} c_{23} Df\, D^2 f\, Df_2 + 360 (2 c_{13}\, A_{113} +$$

$$+ a_3 E_3) f_2 (Df)^2\, Df_1 + 90\, c_{23}^2\, f_2 (D^2 f)^2 + 360\, d_{13}^2 f_1^2 f_2\, (Df)^2 +$$

$$+ 120\, c_{13}^2\, f_3 (Df)^3] \; h^6 + \ldots$$

$$m_4 = f + \ldots + \tfrac{1}{720} \left[a_4^6\, D^6 f + 6\, c_{54} f_1 D^5 f + 30\, d_{44} f_1^2 D^4 f \right. +$$

$$+ 120\, e_{34} f_1^3\, D^3 f + 30\, a_4 c_{44} D^4 f\, Df_1 + 120 (A_{134} + a_4 d_{34}) f_1 D^3 f Df_1 +$$

$$+ 360 (A_{124} + B_{124} + a_4 e_{24}) f_1^2\, D^2 f\, Df_1 + 30\, a_4^4\, c_{14}\, Df\, D^4 f_1 +$$

$$+ 120\, (A_{314} + a_4^3 d_{14}) f_1 Df\, D^3 f_1 + 360\, (A_{214} + B_{214} +$$

$$+ a_4^2\, e_{14}) f_1^2\, Df\, D^2 f_1 + 60\, a_4^2 c_{34} D^3 f D^2 f_1 + 180\, (A_{224} +$$

$$+ a_4^2\, d_{24}) f_1 D^2 f\, D^2 f_1 + 60\, a_4^3\, c_{24} D^2 f\, D^3 f_1 + 360\, (a_4^2\, A_{114} +$$

$$+ a_4\, A_{214}) Df\, Df_1 D^2 f_1 + 120\, c_{14} c_{34} f_2\, Df\, D^3 f + 360\, (c_{14} d_{24} +$$

$$+ c_{24} d_{14} + \mathcal{C}_4)\, f_1 f_2\, Df D^2 f + 360\, a_4 A_{124}\, D^2 f\, (Df_1)^2 +$$

$$+ 720\, (a_4 A_{114} + a_4 B_{114} + H_4) f_1\, Df\, (Df_1)^2 + 180\, a_4^2 c_{14}^2 (Df)^2 D^2 f_2 +$$

$$+ 360\, (2\, a_4 c_{14} d_{14} + E_{14}) f_1 (Df)^2 Df_2 + 360\, a_4 c_{14} c_{24} Df D^2 f Df_2 +$$

$$+ 360\, (2\, c_{14} A_{114} + a_4 E_4) f_2\, (Df)^2 Df_1 + 90\, c_{24}^2 f_2 (D^2 f)^2 +$$

$$+ 360\, (2\, c_{14} e_{14} + 2\, \mathcal{D}_4 + d_{14}^2 + \mathcal{E}_4) f_1^2\, f_2 (Df)^2 +$$

$$\left. + 120\, c_{14}^3\, f_3 (Df)^3 \right] h^6 + \ldots$$

$$m_5 = f + \ldots + \tfrac{1}{720}\,\big[a_5^6 D^6 f + 6\,c_{55} f_1 D^5 f + 30\,d_{45} f_1^2\,D^4 f +$$

$$+ 120\,e_{35} f_1^3\,D^3 f + 360\,g_{25} f_1^4\,D^2 f + 30\,a_5 c_{45} D^4 f D f_1 + 120\,(A_{135} +$$

$$+ a_5 d_{35}) f_1 D^3 f D f_1 + 360\,(\mathcal{A}_{125} + B_{125} + a_5 e_{25}) f_1^2 D^2 f D f_1 +$$

$$+ 720\,(T_5 + \mathcal{B}_5 + \Gamma_5 + a_5 g_{15}) f_1^3 D f D f_1 + 30\,a_5^4 c_{15} D f D^4 f_1 +$$

$$+ 120\,(A_{315} + a_5^3 d_{15}) f_1 D f D^3 f_1 + 360\,(\mathcal{A}_{215} + B_{215} +$$

$$+ a_5^2 e_{15}) f_1^2 D f D^2 f_1 + 60\,a_5^2 c_{35} D^3 f D^2 f_1 + 180\,(A_{225} + a_5^2 d_{25}) f_1 D^2 f D^2 f_1 +$$

$$+ 60\,a_5^3 c_{25} D^2 f D^3 f_1 + 360\,(a_5^2\,A_{115} +$$

$$+ a_5\,A_{215}) D f D f_1 D^2 f_1 + 120\,c_{15} c_{35} f_2 D f D^3 f + 360\,(c_{15} d_{25}) + c_{25} d_{15} +$$

$$+ \mathcal{C}_5) f_1 f_2 D f D^2 f + 360\,a_5 A_{125} D^2 f (D f_1)^2 + 720\,(a_5 \mathcal{A}_{115} + a_5 B_{115} +$$

$$+ H_5) f_1 D f (D f_1)^2 + 180\,a_5^2 c_{15}^2 (D f)^2 D^2 f_2 + 360\,(2\,a_5 c_{15} d_{15} +$$

$$+ E_{15}) f_1 (D f)^2\,D f_2 + 360\,a_5 c_{15} c_{25} D f D^2 f D f_2 + 360\,(2\,c_{15} A_{115} +$$

$$+ a_5 E_5) f_2 (D f)^2 D f_1 + 90\,c_{25}^2 f_2\,(D^2 f)^2 + 360\,(2\,c_{15} e_{15} + 2\,\mathcal{D}_5 + d_{15}^2 +$$

$$+ \mathcal{E}_5) f_1^2 f_2 (D f)^2 + 120\,c_{15}^3 f_3 (D f)^3\big]\,h^6 + \ldots$$

$$m_6 = f + \ldots + \tfrac{1}{720}\,\big[a_6^6\,D^6 f + 6\,c_{56} f_1 D^5 f + 30\,d_{46} f_1^2\,D^4 f +$$

$$+ 120\,e_{36} f_1^3 D^3 f + 360\,g_{26} f_1^4\,D^2 f + 720\,\ell_{16} f_1^5 D f + 30\,a_6 c_{46} D^4 f D f_1 +$$

$$+ 120\,(A_{136} + a_6 d_{36}) f_1 D^3 f D f_1 + 360\,(\mathcal{A}_{126} + B_{126} +$$

$$+ a_6 e_{26}) f_1^2 D^2 f D f_1 + 720\,(T_6 + \mathcal{B}_6 + \Gamma_6 + a_6 g_{16}) f_1^3 D f D f_1 +$$

$$+ 30\,a_6^4 c_{16} D f D^4 f_1 + 120\,(A_{316} + a_6^3 d_{16}) f_1 D f D^3 f_1 + 360\,(\mathcal{A}_{216} +$$

$$+ B_{216} + a_6^2 e_{16}) f_1^2\,D f D^2 f_1 + 60\,a_6^2 c_{36} D^3 f D^2 f_1 + 180\,(A_{226} +$$

$$+ a_6^2 d_{26}) f_1 D^2 f D^2 f_1 + 60\, a_6^3 c_{26} D^2 f D^3 f_1 + 360 (a_6^2 A_{116} +$$

$$+ a_6 A_{216}) Df Df_1 D^2 f_1 + 120\, c_{16} c_{36} f_2 Df D^3 f + 360 (c_{16} d_{26} + c_{26} d_{16} +$$

$$+ \mathcal{C}_6) f_1 f_2 Df D^2 f + 360\, a_6 A_{126}\, D^2 f\, (Df_1)^2 + 720\, (a_6 \mathcal{A}_{116} + a_6 B_{116} +$$

$$+ H_6) f_1 Df\, (Df_1)^2 + 180\, a_6^2 c_{16}^2 (Df)^2 D^2 f_2 + 360\, (2\, a_6 c_{16} d_{16} +$$

$$+ E_{16}) f_1 (Df)^2\, Df_2 + 360\, a_6 c_{16} c_{26} Df D^2 f Df_2 + 360\, (2\, c_{16} A_{116} +$$

$$+ a_6 E_6) f_2 (Df)^2\, Df_1 + 90\, c_{26}^2 f_2 (D^2 f)^2 + 360 (2\, c_{16} e_{16} + 2\, \mathcal{D}_6 +$$

$$+ d_{16}^2 + \mathcal{E}_6) f_1\, f_2 (Df)^2 + 120\, c_{16}^3 f_3\, (Df)^3 \Big]\, h^6 + \ldots$$

$$m_7 = f + \ldots + \frac{1}{720} \Big[a_7^6 D^6 f + 6\, c_{57} f_1 D^5 f + 30\, d_{47} f_1^2 D^4 f +$$

$$+ 120\, e_{37} f_1^3 D^3 f + 360\, g_{27} f_1^4 D^2 f + 720\, \ell_{17} f_1^5 Df + 30\, a_7 c_{47} D^4 f Df_1 +$$

$$+ 120\, (A_{137} + a_7 d_{37}) f_1 D^3 f Df_1 + 360 (\mathcal{A}_{127} + B_{127} + a_7 e_{27}) f_1^2 D^2 f Df_1 +$$

$$+ 720\, (T_7 + \mathcal{B}_7 + \Gamma_7 + a_7 g_{17}) f_1^3 Df Df_1 + 30\, a_7^4 c_{17} Df D^4 f_1 +$$

$$+ 120\, (A_{317} + a_7^3 d_{17}) f_1 Df D^3 f_1 + 360\, (\mathcal{A}_{217} + B_{217} +$$

$$+ a_7^2 e_{17}) f_1^2 Df D^2 f_1 + 60\, a_7^2 c_{37} D^3 f D^2 f_1 + 180\, (A_{227} + a_7^2 d_{27}) f_1 D^2 f D^2 f_1 +$$

$$+ 60\, a_7^3 c_{27} D^2 f D^3 f_1 + 360\, (a_7^2 A_{117} + a_7 A_{217}) Df Df_1 D^2 f_1 +$$

$$+ 120\, c_{17} c_{37} f_2 Df D^3 f + 360 (c_{17} d_{27} + c_{27} d_{17} + \mathcal{C}_7) f_1 f_2 Df D^2 f +$$

$$+ 360\, a_7 A_{127} D^2 f\, (Df_1)^2 + 720\, (a_7 \mathcal{A}_{117} + a\, B_{117} + H_7) f_1 Df (Df_1)^2 +$$

$$+ 180\, a_7^2 c_{17}^2 (Df)^2 D^2 f_2 + 360\, (2\, a_7 c_{17} d_{17} + E_{17}) f_1 (Df)^2\, Df_2 +$$

$$+ 360\, a_7 c_{17} c_{27} Df D^2 f Df_2 + 360\, (2\, c_{17} A_{117} + a_7 E_7) f_2 (Df)^2\, Df_1 +$$

$$+ 90 c_{27}^2\, f_2 (D^2 f)^2 + 360 (2\, c_{17} e_{17} + 2\, \mathcal{D}_7 + d_{17}^2 + \mathcal{E}_7) f_1^2 f_2 (Df)^2 +$$

$$+ 120\, c_{17}^3 f_3 (Df)^3 \Big]\, h^6 + \ldots$$

Parameters in the formulas for $m_2, \ldots, m_7$ are defined by means of following transformations /20/:

$$
\begin{aligned}
c_{ji} &= \sum_{k=2}^{i} a_{k-1}^{j} b_{ik}, \quad i = 2,3,\ldots,7; \quad j = 1,2,\ldots,5 \\
d_{ji} &= \sum_{k=3}^{i} c_{j,k-1} b_{ik}; \quad i = 3,4,\ldots 7; \quad j = 1,2,3,4 \\
A_{rsi} &= \sum_{k=3}^{i} a_{k-1}^{r} c_{s,k-1} b_{ik}; \quad r = 1,2,3,; \quad s = 1,2,3; \quad i = 3,4,5,6,7 \\
E_{1i} &= \sum_{k=3}^{i} a_{k-1} c_{1,k-1}^{2} b_{ik}; \quad i = 3,4,5,6,7 \\
E_{i} &= \sum_{k=3}^{i} c_{1,k-1}^{2} b_{ik}; \quad i = 3,4,5,6,7 \\
e_{ji} &= \sum_{k=4}^{i} d_{j,k-1} b_{ik}; \quad j = 1,2,3; \quad i = 4,5,6,7 \\
\mathcal{A}_{ijk} &= \sum_{t=4}^{k} A_{i,j,t-1} b_{kt}; \quad i = j = 1,2; \quad k = 4,5,6,7 \\
B_{ijk} &= \sum_{t=3}^{k-1} a_{t}^{i} d_{jt} b_{k,t+1}; \quad i = j = 1,2; \quad k = 4,5,6,7 \\
\varphi_{i} &= \sum_{k=4}^{i} c_{1,k-1} c_{2,k-1} b_{ik}; \quad i = 4,5,6,7 \\
H_{i} &= \sum_{k=4}^{i} a_{k-1} A_{1,1,k-1} b_{ik} \\
\mathcal{D}_{i} &= \sum_{k=4}^{i} c_{1,k-1}, d_{1,k-1} b_{ik} \qquad i = 4,5,6,7 \\
\mathcal{E}_{i} &= \sum_{k=4}^{i} E_{k-1} b_{ik}
\end{aligned}
\qquad /20/
$$

$$g_{ji} = \sum_{k=5}^{i} e_{j,k-1}\, b_{ik} \;; \quad j = 1,2;\ i = 5,6,7$$

$$T_i = \sum_{k=5}^{i} \mathcal{E}^{(1)}_{1,k-1}\, b_{ik} \quad \text{where} \quad \mathcal{E}^{(1)}_{1,k-1} = \sum_{t=4}^{k-1} A_{1,1,t-1} b_{k-1,t}$$
$$i = 5,6,7;$$

$$\mathcal{B}_i = \sum_{k=5}^{i} e^{(1)}_{1,k-1} b_{ik} \quad \text{where} \quad e^{(1)}_{1,k-1} = \sum_{t=3}^{k-2} a_t d_{1t} b_{k-1,t+1}$$
$$i = 5,6,7$$

$$\Gamma_i = \sum_{k=5}^{i} a_{k-1} e_{1,k-1} b_{ik}; \quad i = 5,6,7$$

$$\ell_{1i} = \sum_{k=6}^{i} b_{ik} g_{1,k-1} \;; \quad i = 6,7$$

/20/

We determine K from /6/. Using /5/ we obtain

$$\begin{aligned}
\bar{y}(x) = {} & y_o + h\, f(x_o,y_o) + \ldots + h^7 \left[Q_1 D^6 f + Q_2 f_1 D^5 f + Q_3 f_1^2 D^4 f + \right. \\
& + Q_4 f_1^3 D^3 f + Q_5 f_1^4 D^2 f + Q_6 f_1^5 Df + Q_7 D^4 f\, Df_1 + Q_8 f_1 D^3 f\, Df_1 + \\
& + Q_9 f_1^2 D^2 f\, Df_1 + Q_{10} f_1^3\, DfDf_1 + Q_{11} DfD^4 f_1 + Q_{12} f_1\, Df\, D^3 f_1 + \\
& + Q_{13} f_1^2\, DfD^2 f_1 + Q_{14}\, D^3 f D^2 f_1 + Q_{15} f_1\, D^2 f D^2 f_1 + Q_{16}\, D^2 f D^3 f_1 + \\
& + Q_{17}\, DfDf_1 D^2 f_1 + Q_{18} f_2 Df\, D^3 f + Q_{19} f_1 f_2 Df D^2 f + Q_{20} D^2 f\, (Df_1)^2 + \\
& + Q_{21} f_1 Df\, (Df_1)^2 + Q_{22} (Df)^2 D^2 f_2 + Q_{23} f_1 (Df)^2 Df_2 + \\
& + Q_{24}\, DfD^2 f Df_2 + Q_{25} f_2 (Df)^2 Df_1 + Q_{26} f_2 (D^2 f)^2 + \\
& \left. + Q_{27} f_1^2 f_2\, (Df)^2 + Q_{28} f_3 (Df)^3 \right] + O(h^8)
\end{aligned} \qquad /21/$$

where

$$Q_1 = \frac{1}{720}\sum_{i=1}^{7} p_i a_i^6 \; ; \quad Q_2 = \frac{1}{120}\sum_{i=2}^{7} p_i c_{5i} \; ; \quad Q_3 = \frac{1}{24}\sum_{i=3}^{7} p_i d_{4i} \; ;$$

$$Q_4 = \frac{1}{6}\sum_{i=4}^{7} p_i e_{3i} \; ; \quad Q_5 = \frac{1}{2}\sum_{i=5}^{7} p_i g_{2i} \; ; \quad Q_6 = \sum_{i=6}^{7} p_i \ell_{1i}$$

$$Q_7 = \frac{1}{24}\sum_{i=2}^{7} p_i a_i c_{4i} \; ; \quad Q_8 = \frac{1}{6}\sum_{i=3}^{7} p_i \left(A_{13i} + a_i d_{3i}\right)$$

$$Q_9 = \frac{1}{2}\sum_{i=4}^{7} p_i \left(\mathcal{A}_{12i} + B_{12i} + a_i e_{2i}\right)$$

$$Q_{10} = \sum_{i=5}^{7} p_i \left(T_i + \mathcal{B}_i + \Gamma_i + a_i g_{1i}\right)$$

$$Q_{11} = \frac{1}{24}\sum_{i=2}^{7} p_i a_i^4 c_{1i} \; ; \quad Q_{12} = \frac{1}{6}\sum_{i=3}^{7} p_i \left(A_{31i} + a_i^3 d_{1i}\right)$$

$$Q_{13} = \frac{1}{2}\sum_{i=4}^{7} p_i \left(\mathcal{A}_{21i} + B_{21i} + a_i^2 e_{1i}\right) \; ; \quad Q_{14} = \frac{1}{12}\sum_{i=2}^{7} p_i a_i^2 c_{3i} \; ;$$

$$Q_{15} = \frac{1}{4}\sum_{i=3}^{7} p_i \left(A_{22i} + a_i^2 d_{2i}\right) \; ; \quad Q_{16} = \frac{1}{12}\sum_{i=2}^{7} p_i a_i^3 c_{2i} \; ;$$

$$Q_{17} = \frac{1}{2}\sum_{i=3}^{7} p_i \left(a_i^2 A_{11i} + a_i A_{21i}\right) \; ; \quad Q_{18} = \frac{1}{6}\sum_{i=2}^{7} p_i c_{1i} c_{3i} \; ;$$

$$Q_{19} = \frac{1}{2} p_3 \left(c_{13} d_{23} + c_{23} d_{13}\right) + \frac{1}{2}\sum_{i=4}^{7} p_i \left(c_{1i} d_{2i} + c_{2i} d_{1i} + \mathcal{C}_i\right);$$

$$Q_{20} = \frac{1}{2}\sum_{i=3}^{7} p_i a_i A_{12i} \; ; \quad Q_{21} = \sum_{i=4}^{7} p_i \left(a_i \mathcal{A}_{11i} + a_i B_{11i} + H_i\right) \; ;$$

$$Q_{22} = \frac{1}{4}\sum_{i=2}^{7} p_i a_i^2 c_{1i}^2 \; ; \quad Q_{23} = \frac{1}{2}\sum_{i=3}^{7} p_i \left(2\, a_i c_{1i} d_{1i} + E_{1i}\right) \; ;$$

$$Q_{24} = \frac{1}{2}\sum_{i=2}^{7} p_i a_i c_{1i} c_{2i} \; ; \quad Q_{25} = \frac{1}{2}\sum_{i=3}^{7} p_i \left(2\, c_{1i} A_{11i} + a_i E_i\right) \; ;$$

$$Q_{26} = \frac{1}{8} \sum_{i=2}^{7} p_i c_{2i}^2;$$

$$Q_{27} = \frac{1}{2} p_3 d_{13}^2 + \frac{1}{2} \sum_{i=4}^{7} p_i \left(2\, c_{1i} e_{1i} + 2\, \mathcal{D}_i + d_{1i}^2 + \mathcal{E}_i \right);$$

$$Q_{28} = \frac{1}{6} \sum_{i=2}^{7} p_i c_{1i}^3$$

After the substitution /18/ and /21/ into /12/ we obtain

$$\mathcal{J}^*(x_0, x) = \max_{f \in \langle x_0, x \rangle} |R(f)|\, h^7 + O(h^8) \qquad /22/$$

where

$$\begin{aligned}
R(f) = {} & \left(Q_1 - \tfrac{1}{5040}\right) D^6 f + \left(Q_2 - \tfrac{1}{5040}\right) f_1 D^5 f + \left(Q_3 - \tfrac{1}{5040}\right) f_1^2 D^4 f + \\
& + \left(Q_4 - \tfrac{1}{5040}\right) f_1^3 D^3 f + \left(Q_5 - \tfrac{1}{5040}\right) f_1^4 D^2 f + \left(Q_6 - \tfrac{1}{5040}\right) f_1^5\, Df + \\
& + \left(Q_7 - \tfrac{1}{840}\right) D^4 f\, Df_1 + \left(Q_8 - \tfrac{11}{5040}\right) f_1 D^3 f Df_1 + \\
& + \left(Q_9 - \tfrac{1}{336}\right) f_1^2 D^2 f\, Df_1 + \left(Q_{10} - \tfrac{1}{280}\right) f_1^3 Df Df_1 + \left(Q_{11} - \tfrac{1}{336}\right) Df D^4 f_1 + \\
& + \left(Q_{12} - \tfrac{1}{168}\right) f_1 Df D^3 f_1 + \left(Q_{13} - \tfrac{31}{5040}\right) f_1^2 Df\, D^2 f_1 + \\
& + \left(Q_{14} - \tfrac{1}{336}\right) D^3 f\, D^2 f_1 + \left(Q_{15} - \tfrac{5}{1008}\right) f_1 D^2 f\, D^2 f_1 + \\
& + \left(Q_{16} - \tfrac{1}{252}\right) D^2 f\, D^3 f_1 + \left(Q_{17} - \tfrac{9}{560}\right) Df Df_1 D^2 f_1 + \left(Q_{18} - \tfrac{1}{336}\right) f_2 Df D^3 f + \\
& + \left(Q_{19} - \tfrac{1}{112}\right) f_1 f_2 Df D^2 f + \left(Q_{20} - \tfrac{1}{210}\right) D^2 f (Df_1)^2 + \left(Q_{21} - \tfrac{19}{1680}\right) f_1 Df (Df_1)^2 + \\
& + \left(Q_{22} - \tfrac{1}{112}\right) (Df)^2 D^2 f_2 + \left(Q_{23} - \tfrac{5}{336}\right) f_1 (Df)^2 Df_2 + \left(Q_{24} - \tfrac{1}{84}\right) Df D^2 f Df_2 + \\
& + \left(Q_{25} - \tfrac{1}{80}\right) f_2 (Df)^2\, Df_1 + \left(Q_{26} - \tfrac{1}{504}\right) f_2 (D^2 f)^2 + \left(Q_{27} - \tfrac{19}{2520}\right) f_1^2 f_2 (Df)^2 + \\
& + \left(Q_{28} - \tfrac{1}{336}\right) f_3 (Df)^3
\end{aligned}$$

After the evaluation of $Q_1,\dots, Q_{28}$ by means of transformations /20/ we obtain

$$R(\xi) = \frac{11}{37800000} D^6 f - \frac{11}{6300000} f_1 D^5 f - \frac{1}{252000} f_1^2 D^4 f -$$

$$- \frac{1}{31500} f_1^3 D^3 f - \frac{1}{31500} f_1^4 D^2 f - \frac{1}{31500} f_1^5 Df + \frac{1}{252000} D^4 f Df_1 +$$

$$+ \frac{19}{420000} f_1 D^3 f Df_1 + \frac{37}{1260000} f_1^2 D^2 f Df_1 - \frac{1}{14000} f_1^3 Df Df_1 +$$

$$+ \frac{11}{2520000} Df D^4 f_1 - \frac{11}{945000} f_1 Df D^3 f_1 - \frac{19}{420000} f_1^2 Df D^2 f_1 - \frac{17}{2520000} D^3 f D^2 f_1 -$$

$$- \frac{61}{2520000} f_1 D^2 f D^2 f_1 + \frac{11}{1890000} D^2 f D^3 f_1 + \frac{1}{1260000} Df Df_1 D^2 f_1 -$$

$$- \frac{17}{2520000} f_2 Df D^3 f - \frac{139}{7560000} f_1 f_2 Df D^2 f + \frac{1}{63000} D^2 f (Df_1)^2 +$$

$$+ \frac{11}{70000} f_1 Df (Df_1)^2 + \frac{11}{840000} (Df)^2 D^2 f_2 - \frac{11}{1260000} f_1 (Df)^2 Df_2 +$$

$$+ \frac{11}{630000} Df D^2 f Df_2 - \frac{7}{630000} f_2 (Df)^2 Df_1 + \frac{11}{3780000} f_2 (D^2 f)^2 -$$

$$- \frac{23}{756000} f_1^2 f_2 (Df)^2 + \frac{11}{2520000} f_3 (Df)^3 \qquad /23/$$

Being very complicated, formulas /22/, /23/ are inconvenient for the practical using. That is why we have to be satisfied with the estimation of $|\mathcal{Y}(x_0,x)|$. The estimation is done under the assumption that absolute values of the function f and its derivatives up to order 6-th on the interval $\langle x_0,x\rangle$ are not larger than constants suitably chosen. These assumptions are analogical to those appearing in the papers of VEJVODA [3] and JUKL [4] .

I. We assume /according to LOTKIN [5] / that there exist positive constants M,L such that on the interval $\langle x_0,x\rangle$ we have

$$|f| \leq M ; \quad \left| \frac{\partial^n f}{\partial x^{n-i} \partial y^i} \right| \leq \frac{L^n}{M^{i-1}} \qquad /24/$$

for n = 1,2,...,6

i = 0,1,..., n .

In particular, we obtain: $|f_1| \leqq L$; $|f_2| \leqq \frac{L^2}{M}$;

$|f_3| \leqq \frac{L^3}{M^2}$.

The estimation for /13/: $|D^n f| \leqq \sum_{i=0}^{n} \binom{n}{i} \frac{L^n}{M^{i-1}} M^i = 2^n L^n M$

and for /14/: $|D^n f_k| \leqq \sum_{i=0}^{n} \binom{n}{i} \frac{L^{n+k}}{M^{k+i-1}} M^i = 2^n \frac{L^{n+k}}{M^{k-1}}$

Under these assumptions we have, from /23/, the following estimation:

$$|R(\xi)| < 7.10^{-3} L^6 M$$

From /22/ we obtain

$$\mathcal{J}(x_0,x) < 7.10^{-3} L^6 M h^7 + O(h^8) = C_1 \qquad /25/$$

II. The additional assumption L = M in /24/ leads to the special case of /25/ with the only one constant /26/:

$$\mathcal{J}(x_0,x) < 7.10^{-3} (M h)^7 + O(h^8) = C_2 \qquad /26/$$

III. The further special case of LOTKIN assumption is that done by KNICHAL for M=1.

Let L be a positive constant such that $f(x_0,y_0) = 0$; $|f| \leqq 1$;

$|i^f n-i| \leqq L^n$ for $\xi \in \langle x_0,x \rangle$ and $n = 1,2,\ldots,6$.

Then we obtain the estimation

$$\mathcal{J}(x_0,x) < 7.10^{-3} L^6 h^7 + O(h^8) = C_3 \qquad /27/$$

IV. We suppose/according to BIEBERBACH [6] , [7] /that for

$\xi \in \langle x_0, x \rangle$ we have $|f| \leqq M$; $\left| \frac{\partial^n f}{\partial x^{n-i} \partial y^i} \right| = |_{n-i} f_i| \leqq \frac{L}{M^{i-1}}$ /28/

for $n = 1,2,\ldots,6$; $i = 0,1,\ldots,n$

In particular: $|f_1| \leqq L$; $|f_2| \leqq \frac{L}{M}$; $|f_3| \leqq \frac{L}{M^2}$;

Now we have for /13/ and /14/:

$$|D^n f| \leqq 2^n LM \quad ; \quad |D^n f_k| \leqq 2^n \frac{L}{M^k - 1}$$

These assumptions applied to /23/ lead to the following estimation

$$|R(\xi)| \leqq M \left(\frac{11}{590625} L + \frac{428}{590625} L^2 + \frac{1073}{472500} L^3 + \frac{1157}{472500} L^4 + \right.$$

$$\left. \frac{101}{189000} L^5 + \frac{1}{15750} L^6 \right)$$

Denoting the polynom of order 6-th in the brackets by $\varphi(L)$ we have

$$|R(\xi)| \leqq M \varphi(L)$$

Then /22/ leads to

$$\gamma(x_0, x) \leqq M h^7 \varphi(L) + C(h^8) = C_4 \qquad /29/$$

Values of $\varphi(L)$ for some L are given in Table 1.

L	$\varphi(L)$	L	$\varphi(L)$
0.5	0.00065	4.0	1.59115
1.0	0.00606	5.0	4.49453
1.5	0.02650	6.0	10.80791
2.0	0.08145	7.0	23.14513
2.5	0.20340	8.0	45.39402
3.0	0.44238	9.0	83.07769
3.5	0.87115	10.0	143.76154

Table 1

Replacing all coeficients in $\varphi(L)$ by the largest of them, we obtain the simplier but less accurate formula for the estimation

For $L \neq 1$ we have

$$\mathcal{J}(x_0,x) < 0{,}0025 \quad LM \quad \frac{L^6-1}{L-1} \; h^7 + O(h^8) = C_5 \qquad /30/$$

For $L = 1$, we obtain $\mathcal{J}(x_0,x) < 0{,}015 \quad Mh^7 + O(h^8)$

Here is an example for the comparing of the different the error estimations used by VEJVODA [3] in the case RUNGE-KUTTA formula of the 4-th order and by JUKL [4] in the case of NYSTRÖM formula of the 5-th order.

<u>E x a m p l e</u>

$$y' = x^2 + y \,, \quad y(0) = 0$$

The exact solution of this initial problem is:

$$y = 2\,e^x - x^2 - 2x - 2$$

For $x = 0{,}1$; we obtain $y(0{,}1) = 0{,}000\ 341\ 836\ 151\ 295$. Using HUŤA formulas /5/, /10/ gives approximate solution

$$\bar{y}(0{,}1) = 0{,}\ 000\ 341\ 836\ 144\ 687$$

The real error

$$|y(0{,}1) - \bar{y}(0{,}1)| < 7.10^{-12}$$

The error estimation under assumption I.

Conditions of assumption I. are satisfied /see VEJVODA [3] and JUKL [4] /wher $L = 1$, $M = 2$. In this case /25/ leads to

$$|y(0{,}1)| - \bar{y}(0{,}1) < 1{,}4 \,.\, 10^{-9}$$

Huťa formulas /5/, /10/ give about 685 - times better estimation than that by Nyström formula of the 5-th order. Howewer, it is about 200- times worse than the real error.

The estimation under assumption II. Choosing $M = 1{,}26$ /see VEJVODA [3] and JUKL [4] /, we obtain, from /26/

$$|y\ (0,1) - \bar{y}\ (0,1)| < 3{,}6 \cdot 10^{-9}$$

it is a little worse estimation like that in the case I.

The estimation under assumption III. Conditions $f(0,0) = 0$; $|f(x,y)| \leqq 1$ are satisfied on $\langle 0; 0,1 \rangle$.
$|_1 f| \leqq L$; $|f_1| \leqq 1 \leqq L$; $|_2 f| = 2 \leqq L^2$. The choice of $L = \sqrt{2}$; $L \doteq 1{,}42$ leads, by using of /27/ to

$$|y\ (0,1) - \bar{y}\ (0,1)| < 5{,}6.10^{-9}$$

It is the worst estimation in our case.

The error estimation under assumption IV. Choosing $M = 2$, $L = 1$, the conditions will be satisfied. By using /29/, we obtain

$$|y\ (0,1) - \bar{y}(0,1)| < 1{,}4 \cdot 10^{-9}$$

It is the same estimation as in the case I.

When /30/ is used, we obtain

$$|y\ (0,1) - \bar{y}\ (0,1)| < 3. 10^{-9}$$

We conclude that the best estimation was obtained when conditions sub I and sub IV were satisfied.

The error estimation after s-steps

When the solution of equation /1/ with the initial condition /2/ on the interval $\langle x_0, X \rangle$ is sought by using HUŤA's 6-th order formula it is used the recurent method. The interval $\langle x_0, X \rangle$ is divided on N subintervals by the points $x_k = x_0 + kh$ for $k = 1,2,\ldots, N-1$

$$h = \frac{X - x_0}{N}\ , \quad x_0 < x_1 < \ldots < x_N = X.$$

The solution in (r + 1) -st interval is defined as follows:

$$\bar{y}_{r+1}(x) = \bar{y}_r(x_r) + \frac{1}{288}\left(19\,K_{0,r+1} + 75\,K_{3,r+1} + 50\,K_{4,r+1} + \right.$$
$$\left. + 50\,K_{5,r+1} + 75\,K_{6,r+1} + 19\,K_{7,r+1}\right) \qquad /31/$$

where

$$K_{0,r+1} = f\left(x_r,\ \bar{y}_r(x_r)\right)\ (x - x_r)$$

$$K_{1,r+1} = f\left(x_r + \frac{1}{10}(x - x_r);\ \bar{y}_x(x_r) + \frac{K_{0,r+1}}{10}\right)(x - x_r)$$

$$K_{2,r+1} = f\left(x_r\right) + \frac{1}{5}(x - x_r);\ \bar{y}_r(x_r + \frac{-K_{0,r+1}+2K_{1,r+1}}{5}\Big)(x - x_r)$$

$$K_{3,r+1} = f\left(x_r + \frac{x-x_r}{5};\ \bar{y}_r(x_r) + \frac{K_{0,r+1}+4K_{1,r+1}+K_{2,r+1}}{30}\right)(x - x_r)$$

$$K_{4,r+1} = f\left(x_r + \frac{2(x-x_r)}{5};\ \bar{y}_r(x_r) + \frac{4K_{0,r+1}-8K_{1,r+1}-2K_{2,r+1}}{15} + \right.$$
$$\left. + \frac{12K_{3,r+1}}{15}\right)(x - x_r)$$

$$K_{5,r+1} = f\left(x_r + \frac{3(x-x_r)}{5};\ \bar{y}_r(x_r) + \frac{3K_{0,r+1}-4K_{1,r+1}-K_{2,r+1}}{10} + \right.$$
$$\left. + \frac{4K_{3,r+1}+4K_{4,r+1}}{10}\right)(x-x_r)$$

$$K_{6,r+1} = f\Big(x_r + \frac{4(x-x_r)}{5};\ \bar{y}_r(x_r) + \frac{-8K_{0,r+1}+16K_{1,r+1}+4K_{2,r+1}}{15} + \right.$$
$$\left. + \frac{-2K_{3,r+1}-8K_{4,r+1}+10K_{5,r+1}}{15}\right)(x - x_r)$$

$$K_{7,r+1} = f\Big(x_r + (x-x_r);\ \bar{y}_r(x_r) + \frac{169K_{0,r+1}-260K_{1,r+1}-65K_{2,r+1}}{114} + \right.$$
$$\left. + \frac{60K_{3,r+1}+300K_{4,r+1}-180K_{5,r+1}+90K_{6,r+1}}{114}\right)(x - x_r)$$

for r = 0,1,...,N-1 and $\bar{y}_0(x_0) = y(x_0) = y_0$; $\bar{y}_{r+1}(x_r) = \bar{y}_r(x_r)$. We see that the approximate solution defined in such a way is a continuous function in the region B. The function $\bar{y}_{r+1}(x)$ is chosen so that its

Taylor expansion is in accord with the expansion of the function Y_{r+1}, called the local exact solution on the interval $\langle x_r, x_{r+1}\rangle$ up to termsof the 6-th order. The local exact solution is defined so that the equality /1/ is satisfied, but instead of the point (x_0, y_0) this solution goes through the point $(x_r, \bar{y}_r(x_r))$. The error of the local exact solution on the interval $\langle x_r, x_{r+1}\rangle$ is defined as follows

$$\omega_{r+1} = \max_{\xi \in \langle x_r, x_{r+1}\rangle} \left| \bar{y}_{r+1}(\xi) - Y_{r+1}(\xi) \right|$$

As we can see, this is the error of the integration method at one step and hence the above astimations hold.

The error of the approximate solution after s - step is defined as follows

$$\varphi(x_0, x_s) = \max_{\xi \in \langle x_0, x_s\rangle} \left| y(\xi) - \bar{y}(\xi) \right| \qquad /32/$$

Here are two astimations of this error.

I. In the first estimation the idea of BUKOVICS is utilized /see BUKOVICS [8] , VEJVODA [3] /.

We denote by y_{r+1} the function defined on the interval $\langle x_r, x_{r+1}\rangle$ such that its values on this interval is in accord with those of the exact solution y of problem /1/, /2/.

It is obvious, that

$$\left| y_{r+1}(x) - \bar{y}_{r+1}(x) \right| \leq \left| y_{r+1}(x) - \eta_{r+1}(x) \right| + \left| \eta_{r+1}(x) - \bar{y}_{r+1}(x) \right| \qquad /33/$$

where $\eta_{r+1}(x)$ means the approximate solution begining in the point of the exact solution $(x_r, y(x_r))$ for every interval $\langle x_r, x_{r+1}\rangle$. The first component of the right side of /33/ can be estimated by means of results obtained when the error estimation at one step was done.

We now proceed to estimate $\left|\eta_{r+1}(x) - \bar{y}_{r+1}(x)\right|$

$$\eta_{r+1}(x) = y_r(x_r) + \frac{1}{288}\left(19L_{0,r+1} + 75L_{3,r+1} + 50L_{4,r+1} + 50L_{5,r+1} + 75L_{6,r+1} + 19L_{7,r+1}\right) \qquad /34/$$

where

$$L_{0,r+1} = f\left(x_r,\ y_r(x_r)\right)(x-x_r)$$

$$L_{1,r+1} = f\left(x_r + \frac{x-x_r}{10},\ y_r(x_r) + \frac{L_{0,r+1}}{10}\right)(x-x_r)$$

$$L_{2,r+1} = f\left(x_r + \frac{x-x_r}{5},\ y_r(x_r) + \frac{-L_{0,r+1}+2L_{1,r+1}}{5}\right)(x-x_r)$$

$$L_{3,r+1} = f\left(x_r + \frac{x-x_r}{5},\ y_r(x_r) + \frac{L_{0,r+1}+4L_{1,r+1}+L_{2,r+1}}{30}\right)(x-x_r)$$

$$L_{4,r+1} = f\left(x_r + \frac{2(x-x_r)}{5},\ y_r(x_r) + \frac{4L_{0,r+1}-8L_{1,r+1}-2L_{2,r+1}}{15} + \frac{12L_{3,r+1}}{15}\right)(x-x_r)$$

$$L_{5,r+1} = f\left(x_r + \frac{3(x-x_r)}{5},\ y_r(x_r) + \frac{3L_{0,r+1}-4L_{1,r+1}-L_{2,r+1}}{10} + \frac{4L_{3,r+1} + 4L_{4,r+1}}{10}\right)(x-x_r)$$

$$L_{6,r+1} = f\left(x_r + \frac{4(x-x_r)}{5},\ y_r(x_r) + \frac{-8L_{0,r+1}+16L_{1,r+1}+4L_{2,r+1}}{15} + \frac{-2L_{3,r+1}-8L_{4,r+1}+10L_{5,r+1}}{15}\right)(x - x_r)$$

$$L_{7,r+1} = f\left(x_r + \frac{x - x_r}{1},\ y_r(x_r) + \frac{169L_{0,r+1}-260L_{1,r+1}-65L_{2,r+1}}{114} + \frac{60L_{3,r+1}+300L_{4,r+1}-180L_{5,r+1}+90L_{6,r+1}}{114}\right)(x - x_r)$$

By using of the condition /3/ imposed on the function $f(x,y)$ we obtain /35/:

$$|L_{0,r+1}-K_{0,r+1}| \leq Kh\,|y_r(x_r) - \bar{y}_r(x_r)|$$

$$|L_{1,r+1}-K_{1,r+1}| \leq Kh\left(1 + \frac{Kh}{10}\right)|y_r(x_r) - \bar{y}_r(x_r)|$$

$$|L_{2,r+1}-K_{2,r+1}| \leq Kh\left(1 + \frac{-3}{5}Kh + \frac{1}{25}K^2h^2\right)|y_r(x_r) - \bar{y}_r(x_r)|$$

$$|L_{3,r+1}-K_{3,r+1}| \leq Kh\left(1 + \frac{1}{5}Kh + \frac{1}{30}K^2h^2 + \frac{1}{750}K^3h^3\right)|y_r(x_r) - \bar{y}_r(x_r)|$$

$$|L_{4,r+1}-K_{4,r+1}| \leq Kh\left(1 + \frac{26}{15}Kh + \frac{22}{75}K^2h^2 + \frac{4}{125}K^3h^3 + \frac{2}{1875}K^4h^4\right)|y_r(x_r) - \bar{y}_r(x_r)|$$

$$|L_{5,r+1}-K_{5,r+1}| \leq Kh\left(1 + \frac{8}{5}Kh + \frac{131}{150}K^2h^2 + \frac{101}{750}K^3h^3 + \frac{1}{75}K^4h^4 + \frac{4}{9375}K^5h^5\right)|y_r(x_r) - \bar{y}_r(x_r)|$$

$$|L_{6,r+1}-K_{6,r+1}| \leq Kh\left(1 + \frac{16}{5}Kh + \frac{514}{225}K^2h^2 + \frac{848}{1125}K^3h^3 + \frac{602}{5625}K^4h^4 + \frac{266}{28125}K^5h^5 + \frac{8}{28125}K^6h^6\right)|y_r(x_r) - \bar{y}_r(x_r)|$$

$$|L_{7,r+1}-K_{7,r+1}| \leq Kh\left(1 + \frac{562}{57}Kh + \frac{391}{38}K^2h^2 + \frac{759}{190}K^3h^3 + \frac{424}{475}K^4h^4 + \frac{772}{7125}K^5h^5 + \frac{58}{7125}K^6h^6 + \frac{8}{35625}K^7h^7\right)|y_r(x_r) - \bar{y}_r(x_r)|$$

/35/

in this case $h = x - x_r$ indicates variable.

Using of /31/, /34/ and /35/ leads to

$$\max_{\xi \in \langle x_r, x_{r+1}\rangle} |\eta_{r+1}(\xi) - \bar{y}_{r+1}(\xi)| \leq \left(1 + Kh + \frac{203}{96}K^2h^2 + \frac{1283}{864}K^3h^3 + \frac{2113}{4320}K^4h^4 + \frac{241}{2700}K^5h^5 + \frac{523}{54000}K^6h^6 + \frac{11}{18000}K^7h^7 + \frac{1}{67500}K^8h^8\right)|y_r(x_r) - \bar{y}_r(x_r)|$$

The formula /33/ yields the recurrent relation

$$\varphi(x_r, x_{r+1}) \leq q\ \varphi(x_{r-1}, x_r) + C_i \qquad /36/$$

where

$$q = q(Kh) = \left(1 + Kh + \frac{203}{96} K^2h^2 + \frac{1283}{864} K^3h^3 + \frac{2113}{4320} K^4h^4 + \right.$$

$$\left. + \frac{241}{2700} K^5h^5 + \frac{523}{54000} K^6h^6 + \frac{11}{18000} K^7h^7 + \frac{1}{67500} K^8h^8\right)$$

and C_i mean the values obtained by using of /25/, /26/, /27/, /29/ and /30/ for i = 1,2,...,5.

Finally, using of /36/ leads to

$$\varphi(x_o, x_s) \leq C_i \frac{q^s - 1}{q - 1} = C_i \Phi_s(Kh) \qquad /37/$$

$$\Phi_s(Kh) = \frac{q^s - 1}{q - 1}$$

<u>R e m a r k</u>. The constants L,M and the Lipschitz constant K from the preceding paragraph are determined so that the corespodin inequalities are satisfied in the all region B.

In Table 2 values of function Φ_s for some s and hK are given.

	$\Phi_s(hK)$	
s	hK = 0,05	hK = 0,1
1	.10000000' 01	.10000000' 01
2	.20554749' 01	.21226805' 01
3	.31695024' 01	.33830921' 01
4	.43453309' 01	.47981317' 01
5	.55863885' 01	.63867690' 01
6	.68962938' 01	.81703012' 01
7	.82788663' 01	.10172638' 02
8	.97381373' 01	.12420623' 02
9	.11278362' 02	.14944392' 02
10	.12904030' 02	.17777778' 02
20	.35045445' 02	.74328665' 02
30	.73036852' 02	.25421633' 03
40	.13822451' 03	.82643674' 03
50	.25007694' 03	.26466627' 04
60	.44199923' 03	.84367787' 04
70	.77130961' 03	.26855065' 05
80	.13363577' 04	.85443404' 05
90	.23058972' 04	.27181217' 06
100	.39694840' 04	.86464889' 06

Table 2

II. It is possible to show that /see VEJVODA [3] and JUKL [4] /

$$\varphi(x_0,x_s) = \max_{r=0,1,\ldots,s-1} \varphi(x_r,x_{r+1}) \leqq C_i \frac{e^{shK} - 1}{e^{hK} - 1} \qquad /38/$$

REFERENCES

[1] HUŤA A., Une amélioration de la méthode de Runge-Kutta-Nyström pour la résolution numérique des équations defférentielles du premier ordre, Acta F.R.N. Univ. Comen. - Mathematica 1956, 201 - 224.

[2] HUŤA A., Contribution to the numerical solution of differential equations by means of Runge-Kutta formulas with Newton-Cotes numbers weights, Acta F.R.N. Univ. Comen. - Matematica

[3] VEJVODA O., Odhad chyby Runge-Kuttovy formule, Aplikace mat. 2 /1957/, 1-23.

[4] JUKL V., Fehlerabschätzung der Nyström'schen Formel, Acta F.R.N. Univ. Comen. - Matematica XXIV 1970, 81 - 100.

[5] LOTKIN, On the accuracy of Runge-Kutta's Method, Math. Tabl. Oth. Aids. Comp. 5 /1951/, 128 - 133.

[6] BIEBERBACH L., On the remainder of the Runge-Kutta formula in the theory of ordinary differential equations, Z.angew. Math. Phys. Bd. 2 /1951/, 233 - 248.

[7] BIEBERBACH L.. Theorie der Differentialgleichungen /1930/, Berlin.

[8] BUKOVICS E.,Beiträge zur numerischen Integration II, Monatshefte für Math., Bd 57 /1953/, H.4.

Received November 4, 1971, to the publishing January 5, 1972

Author's address: Katedra numerickej matematiky PFUK, Bratislava
Mlýnska dolina, Pavilón matematiky

(ACTA F. R. N. UNIV. COMEN. - MATHEMATICA XXVIII - 1972)

ACTA FACULTATIS RERUM NATURALIUM UNIVERSITATIS COMENIANAE

MATHEMATICA XXVIII - 1972

DECOMPOSITION OF A COMPLETE L-GROUP INTO AN M-PRODUCT OF AN M-ATOMIC AND AN M-NONATOMIC L-SUBGROUPS

TAŤJANA GAVALCOVÁ, Košice

In the paper the following statement is proved: Any complete l-group G can be decomposed into the M-product of its M-atomic and M-nonatomic l-subgroups under the assumption that the l-subgroup M has the following property: the M-carrier belonging to the union of elements from M is not an atom in the lattice of all M-carriers of the l-group G. This is an extension of the result of J.Š.GUREVIČ.

M-disjointness in l-groups and lattices was studied in the papers [2], [3], [7], [8], [9]. The paper [5] of J.Š. GUREVIČ contains the following result (Theorem 3): Any K-group is a direct product of an atomic and a nonatomic K-groups G_a and G_b; by a K-group the author means a complete (in the sense of [4], i.e. the lattice $(G, \vee, \wedge)$ is conditionally complete) divisible l-group. The aim of this paper is to prove an analogous statement on the decomposition of a complete l-group G into the M-product of its M-atomic and M-nonatomic l-subgroups under the assumption that the l-subgroup M satisfies the condition denoted by (C). In our considerations the divisibility of the l-group G is not needed, therefore we get an extension of the result from [5] (even in the case $M = \{0\}$). The assumption that the l-group G is complete is essential as a counterexample of P. JAFFARD [6], p.253) shows.

1. We shall need the following notions.

1.1. The lattice of the M-carriers of the l-group G. Let G be an l-group; the group operation will be written additively. G^+ will denote the positive cone of G, M will be a convex l-subgroup of G. Let us define the relation $\sim$ on G^+ in the following way:

$a, b \in G^+$, $a \sim b \Leftrightarrow a \wedge x \in M$ if and only if $b \wedge x \in M$ for any $x \in G^+$. Then the relation $\sim$ is an equivalence relation on G^+. The classes in the corresponding partition will be denoted $a^\wedge(M) = \{b \in G^+ : b \sim a\}$. The set $a^\wedge(M)$ will be called the M-carrier belonging to the

element a (cf. [2], [7]). It can be easily seen that $a^\wedge(M)$ is a convex subsemigroup and a sublattice in G. The set of all M-carriers of the l-group G will be denoted by $\mathcal{Y}_M(G)$. Let us introduce the partial order on $\mathcal{Y}_M(G)$ in the following way:

$a^\wedge(M) \geqq b^\wedge(M) \Leftrightarrow a \wedge x \in M$ implies $b \wedge x \in M$ for all $x \in G^+$. Then in this partial order $\mathcal{Y}_M(G)$ is a distributive lattice with the least element M; further (cf. [2]),

$$(a \wedge b)^\wedge(M) = a^\wedge(M) \wedge b^\wedge(M),$$

$$(a \vee b)^\wedge(M) = a^\wedge(M) \vee b^\wedge(M) = (a + b)^\wedge(M).$$

The relation $a^\wedge(M) > b^\wedge(M)$ means that $a^\wedge(M) \geqq b^\wedge(M)$ and $a^\wedge(M) \neq b^\wedge(M)$. Let X be an l-subgroup of the l-group G. Let us call the l-subgroup X M-atomic if for any M-carrier $x^\wedge(M) > M$, $0 < x \in X$ there exists an M-carrier $y^\wedge(M)$, $y \in X$, $y^\wedge(M) \leqq x^\wedge(M)$, which is an atom in $\mathcal{Y}_M(G)$. X is called M-nonatomic if it does not contain any element x such that $x^\wedge(M)$ is an atom in $\mathcal{Y}_M(G)$.

1.2. M-product of l-groups. Let M, A, B be convex l-subgroups of the l-group G, $M \subseteq A, B$. G is the M-product of its l-subgroups A, B if the following holds:

(i) (G, +) is generated by subgroups A, B;

(ii) $A \cap B = M$.

In such a case we shall write $G = A \oplus_M B$ and A, B will be called M-summands in G. Further is $B = p(A, M)$ where (cf. [2])

$$P(A, M) = \{y \in G: |y| \wedge |a| \in M \text{ for all } a \in A\}.$$

In the case $M = \{0\}$ we can easily verify that from $G = A \oplus_M B$ it follows that A and B are l-ideals of G; hence the notion of M-product generalizes the notion of direct product (in the case $M = \{0\}$ we obtain the direct product).

Other notions used in the following have the same meaning as in [4].

2. The following considerations will lead to the formulation of the theorem on the decomposition mentioned above.

2.1. Let G be an l-group, let M be its convex l-subgroup. Let $\mathcal{M}_M$ be the set of all atoms in $\mathcal{P}_M(G)$, $\mathcal{N}_M$ be the set of all M-carries which are not greater than or equal to any M-carrier belonging to $\mathcal{M}_M$ (in the partial order by 1.1). Let $\mathcal{M}'_M$ denote the M-complement of the set $\mathcal{N}_M$, i.e. $\mathcal{M}'_M$ is the set of all $a^\wedge(M) \in \mathcal{P}_M(G)$ such that for each M-carrier $b^\wedge(M) \in \mathcal{N}_M$ $a^\wedge(M) \wedge b^\wedge(M) = M$ holds true. Evidently $\mathcal{M}_M \subseteq \mathcal{M}'_M$. Further denote

$$A = \{x \in G^+ : x^\wedge(M) \in \mathcal{M}'_M\},$$
$$B = \{x \in G^+ : x^\wedge(M) \in \mathcal{N}_M\}.$$

Then it is $M^+ \subseteq A, B$. On the other hand if $c \in A \cap B$, $c > 0$, then $c^\wedge(M) \in \mathcal{M}'_M$ and also $c^\wedge(M) \in \mathcal{N}_M$. Thus $c^\wedge(M) \wedge c^\wedge(M) = c^\wedge(M) = M$ and $c \in M^+$. We have proved $A \cap B = M^+$.

2.2. We shall show that B is a subsemigroup in G^+. Let $x, y \in B$ and suppose (by way of contradiction) that there exists an M-carrier $a^\wedge(M) \in \mathcal{M}_M$ such that $a^\wedge(M) \leqq (x + y)^\wedge(M)$. Then we have:

$$x^\wedge(M) \vee (y^\wedge(M) \vee a^\wedge(M)) = x^\wedge(M) \vee y^\wedge(M) \vee a^\wedge(M) =$$
$$= (x + y)^\wedge(M) \vee a^\wedge(M) = (x + y)^\wedge(M) = x^\wedge(M) \vee y^\wedge(M),$$
$$\geqq x^\wedge(M) \wedge (y^\wedge(M) \vee a^\wedge(M)) = (x^\wedge(M) \wedge y^\wedge(M)) \vee (x^\wedge(M) \wedge a^\wedge(M)) =$$
$$= x^\wedge(M) \wedge y^\wedge(M).$$

Because of $x^\wedge(M) \in \mathcal{N}_M$, $a^\wedge(M) \in \mathcal{M}_M \subseteq \mathcal{M}'_M$ we have $x^\wedge(M) \wedge a^\wedge(M) = M$. That means that the M-carriers $y^\wedge(M)$, $y^\wedge(M) \vee a^\wedge(M)$ both are relative complements to the element $x^\wedge(M)$ in the interval $[x^\wedge(M) \wedge y^\wedge(M), x^\wedge(M) \vee y^\wedge(M)]$. From this using the distributivity of the lattice $\mathcal{P}_M(G)$ we obtain

$$y^\wedge(M) = y^\wedge(M) \vee a^\wedge(M),$$

thus $a^\wedge(M) \leqq y^\wedge(M)$ which is a contradiction. We get $x + y \in B$.

2.3. Next we show that B is a convex sublattice in G^+. Let $x, y \in B$; then $(x \vee y)^\wedge(M) \in \mathcal{N}_M$ because $(x \vee y)^\wedge(M) = (x + y)^\wedge(M)$ holds true and that means $x \vee y \in B$. Further let us suppose that there exists an M-carrier $a^\wedge(M) \in \mathcal{M}_M$ such that $a^\wedge(M) \leqq (x \wedge y)^\wedge(M)$. This implies $a^\wedge(M) \leqq x^\wedge(M)$, $a^\wedge(M) \leqq y^\wedge(M)$ which leads to the contradiction with $x, y \in B$. We have proved that B is a sublattice in G^+. If for $x, y \in B$ and $z \in G^+$ $x \leqq z \leqq y$ holds and there exists an M-carrier $a^\wedge(M) \in \mathcal{M}_M$ such that $a^\wedge(M) \leqq z^\wedge(M)$ then $a^\wedge(M) \leqq y^\wedge(M)$ which is a contradiction. Thus $z \in B$ and B is a convex sublattice in G^+.

2.4. Now let us define an M-atomic l-subgroup of G. Put $B_1 = [B]$, i.e. B_1 is a subgroup generated by elements from B, and analogously $A_1 = [A]$. Using 2.2 and 2.3 it can be easily shown that B_1 is a convex l-subgroup in G. From the construction of B_1 we deduce that B_1 is an M-nonatomic subgroup in G. We shall show that A_1 is an M-atomic subgroup. Let $a \in A$, hence $a^\wedge(M) \in \mathcal{M}'_M$ and let $a^\wedge(M) > M$. If there exists no atom in $\mathcal{S}_M(G)$ less than or equal to $a^\wedge(M)$ then it must be $a \in A \cap B$ and hence using 2.1 $a \in M^+$ which is a contradiction. We conclude that A_1 is an M-atomic l-subgroup.

2.5. It will be assumed for the remainder of the section 2 that G is a complete l-group, i.e. the lattice $(G, \vee, \wedge)$ is conditionally complete. Let M be a convex l-subgroup of G satisfying the following condition:

(C) For any system of elements $\{x_\iota\}_{\iota \in I}$ M^+ such that $\bigvee_{\iota \in I} x_\iota$ does exist in G the carrier $\left(\bigvee_{\iota \in I} x_\iota\right)^\wedge(M)$ is not an atom in $\mathcal{S}_M(G)$.

Under this assumption we shall prove that the l-subgroup B_1 is closed in the following sense: B_1 contains all lattice unions of elements from B_1 which exist in G. It suffices to verify that B is closed: let us suppose that B is closed and let $\{x_\iota\}_{\iota \in I}$ be a system of elements from B_1. Let $z = \bigvee_{\iota \in I} x_\iota \in G$. Then $z = z^+ + z^-$ where $z^+ = \left(\bigvee_{\iota \in I} x_\iota\right) \vee 0 = \bigvee_{\iota \in I} (x_\iota \vee 0) \in B$ because of $x_\iota \vee 0 \geq 0$ for all $\iota \in I$ and further

$z^- = \left(\bigvee_{\iota \in I} x_\iota\right) \wedge 0 = \bigvee_{\iota \in I} (x_\iota \wedge 0) \in -B$ because of $x_\iota \wedge 0 \in -B$ for all $\iota \in I$. From this we have $z \in B + (-B) \subseteq B_1$ and B_1 is a closed l-subgroup, too.

2.6. <u>L e m m a .</u> B i s closed if and only if the condition (C) in the l-group G holds true.

<u>P r o o f</u>. If B is closed then (C) is fulfilled because of $M^+ \subseteq B$. On the other hand, let us suppose the condition (C) holding true and let us have a system $\{x_\iota\}_{\iota \in I} \subseteq B$ such that there exists $\bigvee_{\iota \in I} x_\iota = z \in G$. By way of contradiction let us suppose $z \notin B$, i.e. there exists an atom $a^\wedge(M) \in \mathcal{M}_M$ such that $a^\wedge(M) \leq z^\wedge(M)$ (we may assume $a \leq z$). Then

$$x_\iota^\wedge(M) \wedge a^\wedge(M) = (x_\iota \wedge a)^\wedge(M) \leq a^\wedge(M), \; x_\iota^\wedge(M)$$

holds true. But $x_\iota \in B$ and $a^\wedge(M)$ is an atom, thus for each $\iota \in I$

$$(x_\iota \wedge a)^\wedge(M) \neq a^\wedge(M),\ x_\iota^\wedge(M).$$

This implies $(x_\iota \wedge a)^\wedge(M) = M$, hence $(x_\iota \wedge a) \in M$. Using the distributivity law we have

$$\bigvee_{\iota \in I}(x_\iota \wedge a) = (\bigvee_{\iota \in I} x_\iota) \wedge a = z \wedge a = a.$$

But the elements $x_\iota \wedge a$ belong to M and this is the contradiction with (C). We conclude that B is closed and the proof is complete.

2.7. Using results from sections 1 and 2 and Theorem 16, chap. 5, [4] we obtain: B_1 is a direct summand in G and

$$G = B_1 + B_1^{\sigma}$$

holds true, where $B_1^{\sigma} = p(B_1, \{0\})$. In other words G is generated (in the group and lattice sense) by the l-subgroups B_1, B_1^{σ}. Because of $B_1^{\sigma} = p(B_1, \{0\}) \subseteq p(B_1, M)$ G is generated (in the group sense) by the l-subgroups B_1, $p(B_1, M)$. Further it is fulfilled (cf. [2]) $B_1 \cap p(B_1, M) = M$ hence we conclude that G is the M-product of the l-subgroups B_1, $p(B_1, M)$.

2.8. Theorem. Let G be a complete l-group, let M be its convex l-subgroup satisfying the condition (C). Then G can be decomposed into the M-product of its M-atomic and M-nonatomic l-subgroups.

Proof. By 2.7 there is $G = B_1 \oplus_M p(B_1, M)$; it follows from the construction of B_1 that B_1 is an M-nonatomic l-subgroup of the l-group G. Further we have $p(B_1, M) = p(B, M) = [A]$: let us choose an arbitrary $0 \leqq a \in A$ and $0 \leqq b \in B$. Then $a^\wedge(M) \in \mathcal{M}'_M$, $b^\wedge(M) \in \mathcal{N}_M$, hence $(a \wedge b)^\wedge(M) = a^\wedge(M) \wedge b^\wedge(M) = M$ and this implies $a \wedge b \in M$. We have proved $A \subseteq p(B, M)$, but [A] is the least convex l-subgroup containing A and from this we get $[A] \subseteq p(B, M)$. On the other hand, let $x \in p(B, M)$ be an arbitrary element, i.e. $|x| \wedge b \in M$ for all $b \in B$. For M-carriers we have

$$(|x|)^\wedge(M) \wedge b^\wedge(M) = M$$

and this implies $(|x|)^\wedge(M) \in \mathcal{M}'_M$, hence $|x| \in A \subseteq [A]$. Using the convexity of [A] we obtain $x \in [A]$. The inclusion $p(B, M) \subseteq [A]$ has been proved, thus $p(B, M) = [A]$ holds. But $[A] = A_1$ which gives $p(B_1, M) = A_1$. By 2.4 we see that A_1 is the M-atomic l-subgroup. This completes the proof.

2.9. The condition (C) need not be satisfied in any complete l-group. We shall show it on the following example.

E x a m p l e. Let $M(0, 1)$ be the set of all measurable in the sense of Lebesque functions on the interval $(0, 1)$, let $N(0, 1)$ be the subset of measurable functions f on $(0, 1)$ such that their support

$$I_f = \{x \in (0, 1) : f(x) \neq 0\}$$

is a set of the measure zero. Let us define in $M(0, 1)$ the addition + + and the partial order componentwise; in such a manner $M(0, 1)$ is an l-group. It konwn (cf. [1], chap. XV) that the factor l-group $M(0, 1)/N(0, 1)$ is complete. The symbol $\bar{f}$ will mean the class of functions in $M(0, 1)/N(0, 1)$ containing f. Let $H(0, 1)$ denote the set of all classes in $M(0, 1) / N(0, 1)$ corresponding to the bounded functions from $M(0, 1)$. Then $H(0,1)$ is a convex l-subgroup and, therefore, complete. The element $\bar{e}$ where $e(x) = 1$ for all $x \in (0, 1)$ is a strong unit in $H(0, 1)$. Let J_0 be an ideal of measurable subsets of the interval $(0, 1)$ where the basis of J_0 is formed by intervals $\left(\frac{1}{n+1}, 1\right)$ for $n = 1, 2, \ldots$ Let J be a maximal ideal of measurable subsets of the interval $(0, 1)$ containing J_0. Let M be the set of all classes $\bar{f} \in H(0, 1)$ such that there exists $f \in \bar{f}$ with the property $I_f \in J$. Then M is a convex l-subgroup in $H(0, 1)$. The lattice $\mathcal{S}_M(H(0,1))$ contains only two M-carriers: M and $(\bar{e})^{\wedge}(M)$. To prove this, note that $\bar{e} \notin M$ and if $\bar{a} \notin M$, $\bar{a} \wedge \bar{x} \in M$ for a class $\bar{x}$, then there exist $a \in \bar{a}$, $x \in \bar{x}$ such that $I_a \cap I_x \in J$. But J is the maximal ideal and $I_a \notin J$, therefore $I_x \in J$, from which we get $\bar{x} \in M$. Now let f_n be the characteristic function of the interval $\left(\frac{1}{n+1}, 1\right)$ for $n = 1, 2, \ldots$, hence $\bar{f}_n \in M$. We have $\bar{u} = \bigvee_{n=1,2,\ldots} \bar{f}_n \leq \bar{e}$. If $\bar{u} < \bar{e}$ holds true, then $u(x) < e(x)$ for each $x \in P$, where the measure of P is positive. From the σ-additivity of the measure it follows that the measure of $P \cap \left[\frac{1}{n_0+1}, \frac{1}{n_0}\right)$ is positive for some n_0 and therefore $u(x) < f_{n_0}(x)$ on the set of the positive measure. But on the other hand it must be $u(x) \geq f_n(x)$ for each n everywhere but on the set of the zero measure. We conclude $\bigvee \bar{f}_n = \bar{e}$ and that shows that the condition (C) is not satisfied. Neverthless, the decomposition of the group $H(0, 1)$ into the M-product of its M-atomic and M-nonatomic l-subgroups is possible in a trivial way.

2.10. R e m a r k. There is an open question what is the necessary and sufficient condition for the existence of the decomposition into M-product considered in this paper.

References

[1] G. BIRKHOFF, Lattice theory, Providence 1967.
[2] R. D. BYRD, Lattice ordered groups, Dissertation, Tulane University, 1966.
[3] R. D. BYRD, M-polars in lattice ordered groups, Czech. Math. J. 18 (93), 1968, 230 - 239.
[4] L. FUCHS, Partially ordered algebraic systems, London 1963.
[5] JU. Š. GUREVIČ, K elementarnoj teorii strukturno uporjadočennych abelevych grupp i K-linealov, Dokl. Akad. Nauk SSSR, Tom 175, No. 6, 1967, 1213 - 1215.
[6] P. JAFFARD, Contributions à l'étude des groupes ordonnés, J. math. pures et appl. 32 (1953), 203 - 280.
[7] J. JAKUBÍK, M-polars in lattices, Čas. pěst. mat. 95 (1970), 252 - 255.
[8] T. KATRIŇÁK, M-polaren in halbgeordneten Mengen, Čas. pěst. mat. 95 (1970), 416 - 419.
[9] P. G. KONTOROVIČ, K. M. KUTYJEV, K teorii strukturno uporjadočennych grupp, Izv. vyšš. uč. zav., Matematika, 1959, No. 3(10), 112 - 120.

Received June 1, 1971, to the publishing January 5, 1972

Author's address: ul. Kuzmányho 23, Košice

(ACTA F. R. N. UNIV. COMEN. - MATHEMATICA XXVIII - 1972)

ACTA FACULTATIS RERUM NATURALIUM UNIVERSITATIS COMENIANAE

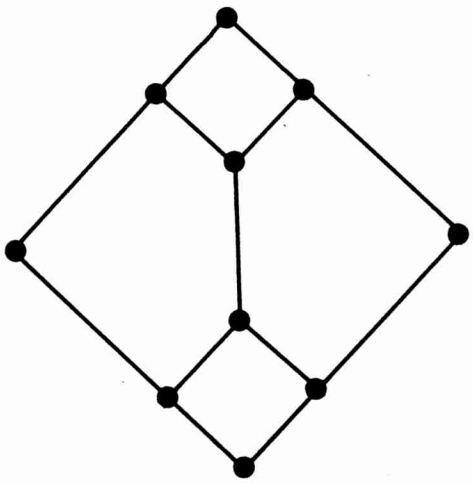
MATHEMATICA XXVIII - 1972

A CHARAKTERIZATION OF p-MODULARITY FOR CONGRUENCE LATTICES OF ALGEBRASS

EVA GEDEONOVÁ, Bratislava

A lattice S is called p-modular iff does not contain a sublattice with diagram of Figure 1. R.MacKenzie in his paper [3] has shown that the lattice S is p-modular if and only if it satisfies the identy

$$(x \cup (y \cap z)) \cap (z \cup (x \cap y)) = (z \cap (x \cup (y \cap z))) \cup (x \cap (z \cup (x \cap y))) .$$



The aim of this paper is to characterize the equational classes A of algebras such that congruences of any algebra of A form a p-modular lattice.

In the paper [2] the following equivalent definition of p-modular lattice was given.

D e f i n i t i o n I. A lattice S is p-modular iff for every $a,b,a_1,b_1 \in S$ the conditions

$$(a_1 \cup b) \cap a = a_1, \quad (a \cup b_1) \cap b = b_1$$

imply $a_1 \cup b_1 = (a_1 \cup b) \cap (a \cup b_1)$.

L e m m a 1. A lattice S is p-modular iff for every $a,b,a_1,b_1 \in S$ the conditions

(1) $(a_1 \cap b) \cup a = a_1, \quad (a \cap b_1) \cup b = b_1$

imply

(2) $a_1 \cap b_1 = (a_1 \cap b) \cup (a \cap b_1)$.

P r o o f. Let S be a p-modular lattice. Let the elements $a,b,a_1,b_1 \in S$ satisfy the conditions (1). Then the elements $a,b,a \cap b_1$, $a_1 \cap b$ satisfy the conditions of Definition I. Hence

$$(a \cap b_1) \cup (a_1 \cap b) = ((a \cap b_1) \cup b) \cap (a \cup (a_1 \cap b)) = b_1 \cap a_1$$

The second implication can be proved dually.

L e m m a 2. In any lattice we have the following identities

$$(a \cup b) \cap c = (((a \cup b) \cap c) \cup b) \cap c,$$
$$(a \cap b) \cup c = (((a \cap b) \cup c) \cap b) \cup c.$$

P r o o f. Clearly $a \cup b \geq ((a \cup b) \cap c) \cup b \geq (a \cup b) \cap c$, hence $(a \cup b) \cap c \geq (((a \cup b) \cap c) \cup b) \cap c \geq (a \cup b) \cap c$.

T h e o r e m 1. A lattice is p-modular if and only if for every $c,d,c',d' \in S$ one of the following identities is satisfied

(3) $((c' \cap d) \cup c) \cap ((c \cap d') \cup d) = (((c' \cap d) \cup c) \cap d) \cup (((c \cap d') \cup d) \cap c)$,

(3') $((c' \cup d) \cap c) \cup ((c \cup d') \cap d) = (((c' \cup d) \cap c) \cup d) \cap (((c \cup d') \cap d) \cup c)$.

Proof. First we suppose that the lattice S satisfies the idenity (3). Then the p-modularity of the lattice S follows from Lemma 1 immediately. Let the lattice S be p-modular. The elements $c, d, (c' \cap d) \cup c$, $(c \cap d') \cup d$ satisfy the conditions (1) of Lemma 1 (see Lemma 2). Hence by (2) the identity (3) follows. Analogously, the lattice S is p-modular if and only if it satisfies the identy (3') (see Definition I).

2.

Let $\mathcal{K}$ be a nontrivial equational class of algebras.

Theorem 2. For an equational class $\mathcal{K}$ of algebras the following two conditions are equivalen.

(i) For every algebra $A \in \mathcal{K}$ the lattice $\Theta(A)$ of all congruenceson A is p-modular.

(ii) Theore exist 6-ary terms $m_0, m_1, \ldots, m_n$ such that for every $\mathcal{K}$-algebra A and every $a,b,c,d,e,f \in A$ there is

(P1) $m_0(a,b,c,d,e,f) = a, \quad (m_n(a,b,c,e,f) = f$

(P2) $m_i(a,a,c,c,a,a) = m_i(a,b,a,a,b,a) = a$, $i = 1,2,\ldots, n-1$

(P3) $m_i(a,b,a,f,b,f) = m_{i+1}(a,b,a,f,b,f)$, i-odd

(P4) $m_i(a,a,c,c,f,f) = m_{i+1}(a,a,c,c,f,f)$, i-even

Proof. First we shall prove that (i) $\longrightarrow$ (ii). Let F be an algebra freely generated by a,b,c,d,e,f over $\mathcal{K}$. Let us define the congruence relations $\mathcal{P}'$, $\mathcal{P}$, ∂', ∂ on the algebra F in the following way: $\mathcal{P}' = \Theta(b,e)$

$$\mathcal{P} = \Theta(b,e) \cup \Theta(a,c) \cup \Theta(d,f)$$

$$\partial' = \Theta(c,d)$$

$$\partial = \Theta(c,d) \cup \Theta(a,b) \cup \Theta(e,f)$$

where $\Theta(x,y)$ denotes the minimal congruence relation induced by $x \equiv y$. Since the lattice $\Theta(F)$ is p-modular we have

(4) $$(((\mathcal{P}' \cup \partial) \cap \mathcal{P}) \cup \partial) \cap (((\mathcal{P} \cup \partial') \cap \partial) \cup \mathcal{P}) \leq ((\mathcal{P}' \cup \partial) \cap \mathcal{P}) \cup ((\mathcal{P} \cup \partial') \cap \partial)$$

(see the identity (3')). From $(b,e) \in \mathcal{P}' \cup \partial$, and $(b,e) \in \mathcal{P}$ it follows that $(b,e) \in (\mathcal{P}' \cup \partial) \cap \mathcal{P}$. Since $(a,b) \in \partial$, $(e,f) \in \partial$ we have

(5) $(a,f) \in (((\mathcal{Y}' \cup \partial) \cap \mathcal{Y}) \cup \partial)$,

Analogously, $(c,d) \in (\mathcal{Y} \cup \partial') \cap \partial$ and $(a,c) \in \mathcal{Y}$, $(d,f) \in \mathcal{Y}$ imply

(6) $(a,f) \in (((\mathcal{Y} \cup \partial') \cap \partial) \cup \mathcal{Y})$.

From (4), (5), (6) it follows

$$(a,f) \in ((\mathcal{Y}' \cup \partial) \cap \mathcal{Y}) \cup ((\mathcal{Y} \cup \partial') \cap \partial)$$

This means that there exists a sequence $u_0, u_1, \ldots, u_n \in F$ such that

(j) $u_0 = a$, $u_n = f$

(jj) $u_i \; (\mathcal{Y}' \cup \partial) \cap \mathcal{Y} \; u_{i+1}$ i-odd

(jjj) $u_i \; (\partial' \cup \mathcal{Y}) \cap \partial \; u_{i+1}$ i-even.

Since the algebra F is freely generated by a,b,c,d,e,f, there exist 6-ary terms $m_0, m_1, \ldots, m_n$ such that

$$u_i = m_i \; (a,b,c,d,e,f), \; i = 0,1,\ldots,n$$

and since the algebra F is freely generated over $\mathcal{K}$ it is sufficient to show that the identities (P1), (P2), (P3), (P4) hold in the algebra F.

Evidently (j) implies (P1). From (jj) it follows $u_i \; (\mathcal{Y}' \cup \partial) \; u_{i+1}$ and $u_i \mathcal{Y} u_{i+1}$. This and the relations $(c,d) \in \mathcal{Y}' \cup \partial$, $(a,b) \in \mathcal{Y}' \cup \partial$, $(a,e) \in \mathcal{Y}' \cup \partial$, $(a,f) \in \mathcal{Y}' \cup \partial$, $(b,e) \in \mathcal{Y}$, $(a,c) \in \mathcal{Y}$, $(d,f) \in \mathcal{Y}$ imply

(7) $m_i(a,a,c,c,a,a) \; \mathcal{Y}' \cup \partial \; m_i(a,b,c,d,e,f) = u_i \; \mathcal{Y}' \cup \partial \; u_{i+1} = m_{i+1}$ $(a,b,c,d,e,f) \; \mathcal{Y}' \cup \partial \; m_{i+1}(a,a,c,c,a,a)$, i-odd

and

(8) $m_i(a,b,a,f,b,f) \; \mathcal{Y} \; m_i \; (a,b,c,d,e,f) = u_i \; \mathcal{Y} \; u_{i+1}$

$= m_{i+1} \; (a,b,c,d,e,f) \; \mathcal{Y} \; m_{i+1} \; (a,b,a,f,b,f)$, i-odd

The algebra $F/\mathcal{Y}' \cup \partial$ is isomorphic to the algebra $F(a,c)$, where $F(a,c)$ denotes the algebra, freely generated by a,c over $\mathcal{K}$. Similarly, the algebra $F/\mathcal{Y}$ is isomorphic to the algebra $F(a,b,f)$ freely generated by a,b,f over $\mathcal{K}$. Hence the relations (7), (8) imply

(9) $m_i \; (a,a,c,c,a,a) = m_{i+1} \; (a,a,c,c,a,a)$, i-odd

(10) $m_i \; (a,b,a,f,b,f) = m_{i+1} \; (a,b,a,f,b,f)$, i-odd.

Analogously, the relation (jjj) implies

(11) $m_i(a,b,a,a,b,a) = m_{i+1}(a,b,a,a,b,a)$, i-even

(12) $m_i(a,a,c,c,f,f) = m_{i+1}(a,a,c,c,f,f)$, i-even,

From (9), (12) and from the relation $a = m_0(a,a,c,c,a,a)$ there follows

$$a = m_i(a,a,c,c,a,a)\ ,\ i = 1,2,\ldots,\ n$$

and from (10),(11) and $a = m_0(a,b,a,a,b,a)$ there follows

$$a = m_i(a,b,a,a,b,a)\ ,\ i = 1,2,\ldots,n.$$

Hence the algebra F satisfies the identity (P2). The algebra F satisfies the identities (P3), (P4) too (see (10),(12)).

Now we prove that (ii) $\longrightarrow$ (i). Let $\mathscr{r}'$, $\mathscr{r}$, ∂', ∂ be congruence relations on a $\mathcal{K}$ -algebra A. Let

(13) $\mathscr{r}' = (\mathscr{r}' \cup \partial) \cap \mathscr{r}$

(14) $\partial' = (\partial' \cup \mathscr{r}) \cap \partial$.

It is sufficient to show that $(\mathscr{r}' \cup \partial) \cap (\mathscr{r} \cup \partial') \leq \mathscr{r}' \cup \partial'$ (see Definition I).

Let us denote for every $i,k \in \{0,1,2,\ldots\} = N$

$\Lambda_i = \partial\mathscr{r}'\partial\mathscr{r}' \ldots \mathscr{r}'\partial$ with $2^{i+1}-1$ factors,

$\Lambda_j = \mathscr{r}\partial'\mathscr{r}\partial' \ldots \partial'\mathscr{r}$ with $2^{k+1}-1$ factors.

Hence $\Lambda_i = \Lambda_{i-1}\mathscr{r}'\Lambda_{i-1}$, $\Lambda_j = \Lambda'_{j-1}\partial'\Lambda'_{j-1}$. The relations Λ_i, Λ'_k are reflexive, symmetrical and compatible with all operations for all $i,k \in N$. The

$$(\mathscr{r}' \cup \partial) \cap (\mathscr{r} \cup \partial') = \bigcup_{i,j \in N} (\Lambda_i \cap \Lambda'_j).$$

Hence it is sufficient to show that

$$\Lambda_i \cap \Lambda_j \leq \mathscr{r}' \cup \partial'$$

for every $i,j \in N$. We shall prove it by induction. From (13), (14) it follows

$$\Lambda_0 \cap \Lambda'_j = \partial \cap \Lambda'_j \leq \partial \cap (\gamma \cup \partial') = \partial' \leq \gamma' \cup \partial',$$

$$\Lambda_i \cap \Lambda'_0 = \Lambda_i \cap \gamma \leq (\gamma' \cup \partial) \cap \gamma = \gamma' \leq \gamma' \cup \partial', \quad i,j \in N.$$

Let $i \neq 0$, $j \neq 0$ and $(a,f) \in \Lambda_i \cap \Lambda'_j$. This implies that there exist elements $b,c,d,e \in A$ such that

(15) $a \; \Lambda_{i-1} b, \; b \; \gamma' e, \; e \; \Lambda_{i-1} f$,

(16) $a \; \Lambda'_{j-1} c, \; c \; \partial' d, \; d \; \Lambda'_{j-1} f$.

Since $\partial' \leq \partial \leq \Lambda_{i-1}$ by (14) and since $\gamma' \leq \gamma \leq \Lambda'_{j-1}$ by (15) we have from the relation (16), (15)

(17) $c \; \Lambda_{i-1} d$

(18) $b \; \Lambda'_{j-1} e$.

For every $k,k = 0,1,\ldots, n$, we have by (15), (17), (P1), (P2)

$$u_k \; \Lambda_{i-1} m_k (b,b,c,c,e,e) \; \gamma' \; m_k(b,b,c,c,b,b) = b = m_k(b,b,b,b,b,b) \gamma'$$
$$\gamma' m_k (b,b,b,e,b,e) \; \Lambda_{i-1} m_k(a,b,a,f,b,f).$$

Therefore $u_k \; \Lambda_{i-1} \gamma' \Lambda_{i-1} m_k(a,b,a,f,b,f)$, hence

(19) $u_k \; \Lambda_i m_k(a,b,a,f,b,f)$,

For every $k,k = 0,1,\ldots,n$, we have by (16), (18), (P1), (P2)

$$u_k \; \Lambda'_{j-1} m_k(c,b,c,d,b,d) \, \partial' \; m_k(c,b,c,c,b,c) = c = m_k(c,c,c,c,c,c) \partial'$$
$$\partial' m_k (c,c,c,d,d) \; \Lambda'_{j-1} m_k(a,a,c,c,f,f). \text{ Hence}$$

(20) $u_k \; \Lambda'_j m_k (a,a,c,c,f,f)$.

For k odd we have by (16), (18), (P3)

$$u_k \Lambda'_{j-1} m_k(a,b,a,f,b,f) = m_{k+1}(a,b,a,f,b,f) \; \Lambda'_{j-1} u_{k+1}.$$

From this and from (19) it follows

$$u_k \Lambda_i \cap \Lambda'_{j-1} m_k(a,b,a,f,b,f) = m_{k+1}(a,b,a,f,b,f) \Lambda_i \cap \Lambda_{j-1} u_{k+1}.$$

By the induction assumption we have $\Lambda_i \cap \Lambda'_{j-1} \leqq \mathcal{P}' \cup \mathcal{P}'$, hence $u_k \; \mathcal{P}' \cup \mathcal{P}' \; u_{k+1}$ for k odd. For k even we have by (15), (17), (P4)

$$u_k \; \Lambda_{i-1} m_k(a,a,c,c,f,f) = m_{k+1}(a,a,c,c,f,f) \; \Lambda_{i-1} u_{k+1}.$$

This and (20) imply

$$u_k \Lambda_{i-1} \cap \Lambda'_j m_k \; (a,a,c,c,f,f) = m_{k+1}(a,a,c,c,f,f) \; \Lambda_{i-1} \cap \Lambda'_j u_{k+1}.$$

By the induction assumption we have $u_k \; \mathcal{P}' \cup \mathcal{P}' u_{k+1}$ for k even. Therefore $a = u_0 \; \mathcal{P}' \cup \mathcal{P}' \; u_1 \; \mathcal{P}' \cup \mathcal{P}' \; . \; . \; . \; \mathcal{P}' \cup \mathcal{P}' \; u_n = f$ hence $(a,f) \in \mathcal{P}' \cup \mathcal{P}'$ as desired.

3.

We shall call the equational classes of algebras having permutable congruence relations as permutable equational classes of algebras.

Definition. An equational class of algebras is n - - p-modular ($n \in N$) iff there exists a sequence of n + 1 6-terms which satisfy the condition (ii) of Theorem 2.

Theorem 3. An equtional class of algebras is permutable if and only if it is 2 - p-modular.

Proof. If the equational class $\mathcal{K}$ of algebras is permutable then by [4] there exists a ternary term p such that $p(a,a,c) = c, p(a,c,c) = a$. Let us define $m_0(a,b,c,d,e,f) = a$, $m_2(a,b,c,d,e,f) = f$ and $m_1(a,b,c,d,e,f) = p(p(a,b,e), p(c,d,f), p(b,e,f))$. The condition (P1) is satisfied by the definition. Since

$$m_1(a,a,c,c,a,a) = p(a,a,a) = a, m_1(a,b,a,a,b,a) = p(a,a,a) = a$$

the condition (P2) is satisfied too. Since

$$m_1(a,b,a,f,b,f) = p(p(a,b,b), p(a,f,f), p(b,b,f) = p(a,a,f) = f$$

$$m_1(a,a,c,c,f,f) = p(p(a,a,f), p(c,c,f), p(a,f,f)) = p(f,f,a) = a$$

the equational class $\mathcal{K}$ of algebras is 2 - p-modular.

If $\mathcal{K}$ is 2 - p-modular then there exist 6-ary terms m_0, m_1, m_2 which satisfy the identities (P1), (P2), (P3), (P4). Let us define

$$p(a,b,c) = m_1(a,a,b,c,b,c).$$

Then $p(a,a,b) = m_1(a,a,a,b,a,b) = m_2(a,a,a,b,a,b) = b$

$p(a,b,b) = m_1(a,a,b,b,b,b) = m_0(a,a,b,b,b,b) = a$

Hence by Malćevs theorem (see [4]) the equational class of algebras is permutable.

Remark. If we use the Definition of the n-modularity of the paper [1] it is easy to prove that if an equational class of algebras is n-modular it is n - p-modular.

References

[1] A.DAY, A charakterization of modularity for congruence lattices of algebras; Canad. Math. Bull. 12 (1969), 167-173

[2] E.GEDEONOVÁ, Jordan-Hölder theorem for lines (to appear in Mat. časop.)

[3] R.MACKENZIE, Equational Bases and Non-modular Lattice Varieties (to appear)

[4] A.J.MALĆEV, On the General Theory of Algebraic Systems, Mat.Sb. 35(77), (1954), 3-20.

Received August 5, 1971, to the publishing January 5, 1972

Author's address: Katedra algebry a teórie čísel, Bratislava, Šmeralova 2

Štrauch O., Minimálne pokrytie uzavretého intervalu 1

Kodnár R., Bemerkungen zu einem System von nichtlinearen Gleichungen 17

Huťa A., Contribution to the numerical Solution of Differential Equations by means of Runge-Kutta Formulas with Newton - Cotes numbers weights 51

Berecová A., An error estimation a Huťa formula of the Runge-Kutta type of the sixth order 67

Gavalcová T., Decomposition of a complete L-group into an M-product of an M-atomic and M-nonatomic l-subgroups 91

Gedeonová E., A charakterization of p-modularity for congruence lattices of algebras . . 99