TW

stichting mathematisch centrum

MC

AFDELING TOEGEPASTE WISKUNDE TN 65/72 JUNI

R. PLOEGER
DIRECTE METHODES

TW

2e boerhaavestraat 49 amsterdam

Printed at the Mathematical Centre, 49, 2e Boerhaavestraat, Amsterdam.

The Mathematical Centre, founded the 11-th of February 1946, is a non-profit institution aiming at the promotion of pure mathematics and its applications. It is sponsored by the Netherlands Government through the Netherlands Organization for the Advancement of Pure Research (Z.W.O), by the Municipality of Amsterdam, by the University of Amsterdam, by the Free University at Amsterdam, and by industries.

Inhoud

0. Inleiding

We zullen ons bezighouden met het toepassen van directe methodes afkomstig uit de variatierekening voor het oplossen van differentiaalvergelijkingen; ook zullen we deze methodes gebruiken om eigenwaarden te bepalen van differentiaaloperatoren. Dit soort problemen komt veel voor in de mechanica (trillingen van snaren, staven), in de elasticiteitstheorie (buiging van platen), quantummechanica (grondtoestand van atomen), en bij het numeriek oplossen op zichzelf van deze problemen: b.v. bij het oplossen van randwaardeproblemen wordt de differentiaaloperator gediscretiseerd, d.w.z. door differenties weergeven ($\frac{du}{dx} \approx \frac{u(x+h)-u(x-h)}{2h}$, voor kleine $h > 0$), en vervolgens wordt het probleem iteratief opgelost. Daarbij is kennis nodig van de ligging van de eigenwaarden van de gediscretiseerde operator, b.v. bij de iteratiemethode van Richardson.
De variatierekening brengt dergelijke problemen in verband met het minimaliseren van functionalen. Dit minimaliseren gebeurt door de directe methodes.
We zullen nu het probleem abstract formuleren.
Zij H een complexe separabele Hilbertruimte, en A een lineaire operator die gedefinieerd is op een dicht gebied $D(A)$ in H.
Dan beschouwen we in H voor een willekeurige $f \in H$ het probleem

(0.1) $\quad Au = f.$

Dit probleem heeft niet altijd een unieke oplossing voor een willekeurige lineaire operator A. Daarvoor moeten aan A bepaalde eisen opgelegd worden, b.v. inverteerbaarheid, d.w.z. A heeft een begrensde inverse.
In H definiëren we de volgende functionalen:

(0.2a) $\quad F_1(d) \overset{\text{def}}{=} (Au,u) - (u,f) - (f,u) \quad ,\forall u \in D(A),$ en

(0.2b) $\quad F_2(u) \overset{\text{def}}{=} ||Au-f||^2 = (Au,Au) - (Au,f) - (f,Au) + ||f||^2, \forall u \in D(A),$

met f het rechterlid uit (0.1).
Er geldt nu:
Als A een positief definiete operator is, d.w.z. er bestaat een $\alpha > 0$ zodat $\forall u \in D(A) \quad (Au,u) \geq \alpha \, ||u||^2$, dan is het oplossen van (0.1)

equivalent met het minimaliseren van (0.2a) of (0.2b); als A een begrensde inverse heeft, dan is het oplossen van (0.1) equivalent met het bepalen van het minimum van (0.2b), zie Mikhlin [6] en [7].
Dit minimum wordt als volgt bepaald:
in H kiezen we een volledig stelsel $\{\phi_k\}$, de coördinaatfuncties;
zij $u_n = \sum_{k=1}^{n} \alpha_k\phi_k$, waarin $\alpha_k (k=1,2,\ldots,n)$ bepaald worden uit de voorwaarde dat:
of: $F_1(u)$ is minimaal in de deelruimte L opgespannen door $\phi_1, \ldots, \phi_n$ (methode van Ritz),
of: $F_2(u)$ is minimaal in de deelruimte L opgespannen door $\phi_1, \ldots, \phi_n$ (methode van de kleinste kwadraten).
Deze voorwaarden leiden tot de respectievelijke stelsels

(0.3a) $$\sum_{k=1}^{n} (A\phi_k,\phi_l)\alpha_k = (f,\phi_l) \quad , \quad l = 1, 2, \ldots, n \qquad \text{(Ritz)};$$

(0.3b) $$\sum_{k=1}^{n} (A\phi_k,A\phi_l)\alpha_k = (f,A\phi_l), \quad l = 1, 2, \ldots, n \quad \text{(kleinste kwadraten)}.$$

Mikhlin heeft in [6] en [7] afgeleid dat geldt $\lim_{n\to\infty} u_n = u_0$, met: u_0 minimaliseert $F_1(u)$ resp. $F_2(u)$.
We definiëren een algemenere functionaal $F_3(u)$ die tot een generalisatie van de methode van Ritz leidt; Mikhlin heeft in [6] opgemerkt dat het minimaliseren van de functionaal $F(u) = \Phi(u,u) - l(u) - \overline{l(u)}$, waarin Φ een kwadratische functionaal en l een lineaire functionaal is, op analoge wijze geschiedt als het minimaliseren van de functionalen $F_1(u)$ en $F_2(u)$.
We kiezen nu $\Phi(u,u) \overset{\text{def}}{=} (Au,Ku)$ en $l(u) \overset{\text{def}}{=} (f,Ku)$, waarin K een lineaire operator is die aan bepaalde voorwaarden voldoet, zodat

(0.2c) $$F_3(u) \overset{\text{def}}{=} (Au,Ku) - (Ku,f) - (f,Ku) \quad , \quad \text{voor alle } u \in D(A).$$

Bovenstaande keuze van Φ en l is afkomstig van Petryshyn (zie [9]) die de analogie heeft uitgewerkt en heeft bewezen dat het minimaliseren van (0.2c) equivalent is met het oplossen van (0.1).
We merken op, dat de methode van Ritz en de methode van de kleinste kwadraten speciale gevallen zijn van deze generalisatie (K=I en K=A); in [9] wordt voor een speciale keuze van K afgeleid, dat ook de methode van de momenten en de methode van Galerkin speciale gevallen van de gegeneraliseerde methode zijn: we zullen deze laatste methodes niet behandelen.

In hoofdstuk 1 schetsen we de achtergronden van de gegeneraliseerde methode. Het zou te ver voeren om alle bewijzen van de benodigde stellingen te formuleren; daarvoor verwijzen we naar [9].
Hoofdstuk 2 behandelt de methode zelf; ook de convergentie van de methode komt aan de orde. Afgeleid wordt dat het minimaliseren van (0.2c) tot het volgende algebraïsche stelsel leidt:

$$(0.3c) \qquad \sum_{k=1}^{n} (A\phi_k, K\phi_l)\alpha_k = (f, K\phi_l) \quad , \; l = 1, 2, \ldots, n \quad \text{(gegeneraliseerde methode van Ritz).}$$

In de laatste paragraaf van dit hoofdstuk wordt een tweetal voorbeelden gegeven. Het derde hoofdstuk bevat de behandeling van het vraagstuk van het bepalen van de eigenwaarden van een lineaire operator A. Als A een zelfgeadjungeerde, positief definiete operator is in een reële separabele Hilbertruimte H, dan is het bepalen van de eigenwaarden van A equivalent met het minimaliseren van de functionaal

$$(0.4a) \qquad G_1(u) \overset{\text{def}}{=} \frac{(Au,u)}{(u,u)} \quad , \quad \text{voor alle } u \in D(A),$$

onder bepaalde nevenvoorwaarden (zie [6] of [7]). Het minimaliseren ervan wordt op analoge wijze uitgevoerd als van $F_i(u)$ $(i=1,2,3)$.
Ook nu definiëren we een algemenere functionaal $G_2(u)$, waarvan bewezen wordt dat het minimaliseren ervan equivalent is met het bepalen van de eigenwaarden van A:

$$(0.4b) \qquad G_2(u) \overset{\text{def}}{=} \frac{(Au,Ku)}{(u,Ku)} \quad , \quad \text{voor alle } u \in D(A).$$

Het minimaliseren van (0.4b) wordt uitgevoerd met de gegeneraliseerde methode van Ritz. De in dit hoofdstuk geformuleerde stellingen zijn analogieën van stellingen uit [6] en [7] voor het geval K = I, en ook de bewijzen ervan verlopen op dezelfde manier. Na de theorie behandeld te hebben, geven we weer enkele voorbeelden.
We merken op, dat bovenstaande methodes leiden tot bovengrenzen van de eigenwaarden; daarom is tot slot een paragraaf gewijd aan het bepalen van ondergrenzen door een methode die voortkomt uit de theorie van de integraalvergelijkingen. Eerst behandelen we in het kort de theorie (voor de bewijzen van de gebruikte stellingen zie [5]), en vervolgens wordt een voorbeeld gegeven.

1. Theoretische achtergronden

1.1. K-positief definiete operatoren

Zij H een separabele, complexe Hilbertruimte; laat A een lineaire, niet noodzakelijk begrensde operator zijn met definitiegebied $D(A)$ dat dicht in H ligt.

We beschouwen voor een willekeurige $f \in H$ het probleem $Au = f$.

Deze vergelijking heeft niet altijd een oplossing; we zullen daarom aan A enkele voorwaarden moeten opleggen. Daartoe voeren we een lineaire operator K in, gedefiniëerd op een dicht gebied $D(K)$ in H. K heeft de volgende eigenschappen:

1^0. K is gesloten: als een rij $\{x_n\} \subset D(K)$ naar x_0 convergeert in $D(K)$ en als de rij $\{Kx_n\}$ naar y_0 convergeert, dan geldt dat $x_0 \in D(K)$ en $Kx_0 = y_0$;

2^0. $D(K) \supseteq D(A)$;

3^0. de beeldruimte $R(K)$ ligt dicht in H;

4^0. K heeft een begrensde inverse.

We noemen nu de operator A een K-positief definiete operator(K-p.d.operator), als er een constante $\alpha > 0$ bestaat zô dat voor iedere $u \in D(A)$ de relatie geldt

$$(1.1) \qquad (Au,Ku) \geq \alpha ||Ku||^2.$$

Als $K = I$, dan is A blijkbaar positief definiet, en als $K = A$, dan is A een inverteerbare operator. We geven een voorbeeld van een niet-triviale K-p.d.operator:

voorbeeld:

Zij $H = L_2(0,1)$; $A = \dfrac{-d^3}{dx^3}$: $D(A) = C^3(0,1)$ ligt dicht in H.

Beschouw het probleem $Au = f$ met randvoorwaarden $u = 0$ voor $x = 0$ en $u'(0) = u'(1) = 0$, voor een gegeven $f \in H$. Neem nu $K = \dfrac{d}{dx}$, dan voldoet K aan de voorwaarden $1^0 - 4^0$. We leiden af dat er een constante $\alpha > 0$ bestaat, waarvoor aan (1.1) voldaan is:

$$\left(\frac{du}{dx}\right)^2 = \left[\int_0^x \left(\frac{d^2u}{dt^2}\right)dt\right]^2 \leq x \int_0^x \left(\frac{d^2u}{dt^2}\right)^2 dt \leq x \int_0^1 \left(\frac{d^2u}{dt^2}\right)^2 dt \Rightarrow \left(\frac{du}{dx},\frac{du}{dx}\right) \leq 2 \left(\frac{d^2u}{dx^2},\frac{d^2u}{dx^2}\right) .$$

Nu is $(Au,Ku) = -\int_0^1 \frac{d^3u}{dx^3}\frac{du}{dx}\,dx = \int_0^1 \left(\frac{d^2u}{dx^2}\right)^2 dx = \left(\frac{d^2u}{dx^2}, \frac{d^2u}{dx^2}\right)$ en $||Ku||^2 = \left(\frac{du}{dx}, \frac{du}{dx}\right)$.

We hebben dus een $\alpha > 0$, nl. $\alpha = \frac{1}{2}$, gevonden waarvoor (1.1) geldt. Uit dit voorbeeld blijkt wel, dat iedere differentiaaloperator van oneven orde K-p.d. is, mits hij op een compact gebied gedefinieerd is.
We kunnen voor de operator A enkele eigenschappen afleiden.

<u>eigenschap 1</u>:

Als A K-p.d. is, dan heeft A een begrensde inverse.

<u>bewijs</u>:

omdat K een begrensde inverse heeft, bestaat $\beta > 0$ met $||Ku|| \geq \beta||u||$....(1)

$||Au||.||Ku|| \geq |(Au,Ku)| \geq \alpha\ ||Ku||^2 \Longrightarrow ||Au|| \geq \alpha\ ||Ku||$...............(2)

Uit (1) en (2) volgt $||Au|| \geq \alpha\beta||u||$. Dus A heeft een begrensde inverse.

<u>eigenschap 2 (symmetrie)</u>:

Als A K-p.d. is, dan geldt voor alle $u,v \in D(A)$: $(Au,Kv) = (Ku,Av)$.

<u>bewijs</u>:

$(A(u+v), K(u+v)) - (A(u-v), K(u-v)) + i(A(u+iv),K(u+iv)) - i(A(u-iv), K(u-iv)) =$

$= 4(Au,Kv)$;
een analoge uitdrukking geldt voor A en K verwisseld.
Nu is $(Aw,Kw) = (Kw,Aw)$, omdat (Aw,Kw) reëel is. Dit impliceert $(Au,Kv) = (Ku,Av)$.

<u>eigenschap 3</u>:

Als A K-p.d. is, dan geldt voor alle $u,v \in D(A)$ $\quad |(Au,Kv)|^2 \leq (Au,Ku).(Av,Kv)$.

<u>bewijs</u>:

Zij λ een willekeurig reëel getal; beschouw voor iedere $u,v \in D(A)$
$w = u + \lambda\ (Au,Kv)\ v \in D(A)$. Nu geldt:

$0 < (Aw,Kw) = (Au,Ku) + 2\lambda\ |(Au,Kv)|^2 + \lambda^2\ |(Au,Kv)|^2\ (Av,Kv)$. Hieruit volgt:

$4(Au,Ku).(Av,Kv)\ .\ |(Au,Kv)|^2 - 4\ |(Au,Kv)|^4 \geq 0.$

eigenschap 4:

Als A K-p.d. is, dan kan A uitgebreid worden tot een gesloten operator.

bewijs:

laat $\{u_n\} \subset D(A)$ met $u_n \to 0$ en $Au_n \to f$ in H voor $n \to \infty$; bij het bewijs van eig. 1 hebben we al uit (1.1) afgeleid dat $||Au|| \geq \alpha ||Ku||$ voor alle $u \in D(A)$ $\Rightarrow$ $\{Ku_n\}$ convergeert in H. Omdat K gesloten is, geldt $Ku_n \to 0$ voor $n \to \infty$. Hieruit volgt voor alle $v \in D(A)$:

$(f,Kv) = \lim_{n\to\infty} (Au_n,Kv) = \lim_{n\to\infty} (Ku_n,Av) = 0$; dus $f \equiv 0$, want R(K) ligt dicht in H, en K heeft een begrensde inverse.

De symmetrie-eigenschap van het inprodukt (Au,Kv) kan niet meer bewezen worden, als we H reëel nemen; in dat geval moet eigenschap 2 geëist worden als voorwaarde waaraan de operator A moet voldoen.
Ten aanzien van het probleem Au = f, voor willekeurige $f \in H$, geldt voor een K-p.d. operator de volgende stelling:

stelling 1.1

Als A een K-p.d.operator is en D(A) = D(K), dan bestaat een constante $\theta > 0$ zodat voor alle $u \in D(K)$

(1.2) $\quad ||Au|| \leq \theta\, ||Ku||$.
Bovendien is A gesloten, R(A) = H, en de vergelijking Au = f heeft voor willekeurige $f \in H$ een en slechts een oplossing.

bewijs zie [9] en [10].

1.2. Variatieprincipe

We voeren de volgende functionaal in:

$F(u) \overset{def}{=} (Au,Ku) - (Ku,f) - (f,Ku)$, voor een willekeurige $f \in H$, en voor alle $u \in D(A)$.

Deze functionaal hangt samen met de vergelijking Au = f.

stelling 1.2

Als A K-p.d. is, dan is een noodzakelijke en voldoende voorwaarde dat u_0 een oplossing is van Au = f, dat F(u) minimaal is voor $u = u_0$.

<u>bewijs</u>:

Stel $Au_0 = f$. Dan is $F(u) = (Au,Ku) - (Ku,Au_0) - (Au_0,Ku)$. Wegens eigenschap 2 geldt:

$$F(u) = (A(u-u_0), K(u-u_0)) - (Au_0,Ku_0).$$

Hieruit zien we onmiddellijk, dat $F(u)$ minimaal is voor $u = u_0$.

Stel $F(u)$ is minimaal voor $u = u_0$; zij $v \in D(A)$ willekeurig en t een reëel getal. Dan geldt:

$$\frac{d}{dt} F(u_0+tv) = 0 \text{ voor } t = 0, \text{ zodat } (Au_0,Kv) + (Av,Ku_0) - (Kv,f) - (f,Kv) = 0.$$

Wegens de symmetrie (eig.2) geldt nu $(Au_0-f,Kv) + (Kv,Au_0-f) = 0$, zodat $\mathrm{Re}(Au_0-f, Kv) = 0$ en met vervanging van v door iv: $\mathrm{Im}\,(Au_0-f,Kv) = 0$.
Voor alle $v \in D(A)$ geldt dus $(Au_0-f,Kv) = 0$.
Omdat $R(K)$ dicht in H ligt, en K een begrensde inverse heeft, geldt dat $Au_0 = f$.

In het geval dat $K = I$ valt dit variatieprincipe samen met het gewone variatieprincipe, n.l. het minimaliseren van $F(u) = (Au,u) - (u,f) - (f,u)$; voor $K = A$ valt de methode samen met de methode van de kleinste kwadraten die $||Au-f||^2$ minimaliseert: $F(u) = ||Au-f||^2 - ||f||^2$.

Het kan voorkomen dat $F(u)$ het minimum bereikt voor $u = u_0$, met $u_0 \notin D(A)$, $u_0 \in H$.
Daarom breiden we het definitiegebied $D(A)$ uit tot een Hilbertruimte H_K door de invoering van een nieuw inprodukt:

(1.3) $[u,v] \overset{\text{def}}{=} (Au,Kv)$, voor alle u, $v \in D(A)$.

Hiermee ligt tevens een norm vast, nl. $|u| = [u,u]^{1/2}$, voor alle $u \in D(A)$.
Vervolgens maken we $D(A)$ volledig m.b.t. deze norm, en vormen zo de Hilbertruimte H_K.
Uit $(Au,Ku) \geq \alpha||Ku||^2$ en $||Ku|| \geq \beta||u||$, voor een $\alpha, \beta > 0$, leiden we de relaties af:

(1.4) $|u| \geq \gamma_1 ||Ku||$, waarin $\gamma_1 = \sqrt{\alpha} > 0$, en $u \in D(A)$;

(1.5) $|u| \geq \gamma_2 ||u||$, waarin $\gamma_2 = \beta\sqrt{\alpha} > 0$, en $u \in D(A)$.

De Hilbertruimte H_K heeft een aantal eigenschappen:

1^0. D(A) ligt dicht in H_K;

2^0. H_K is een deelruimte van H door identificatie van de elementen uit H_K met elementen in H; bewijs zie [9].

3^0. K kan uitgebreid worden tot een begrensde lineaire operator van H_K naar H;
dit volgt uit ongelijkheid (1.4).

4^0. de ongelijkheden (1.4) en (1.5) zijn geldig voor alle $u \in H_K$.

In H_K heeft het minimaliseringsprobleem van F(u) een oplossing op grond van de stelling:

stelling 1.3

Als A K-p.d. is, dan heeft het probleem van het minimaliseren van F(u) een en slechts een oplossing u_0 in H_K voor willekeurige $f \in H$; als $\{\phi_k\}$ een volledig orthonormaal stelsel is* in H_K, dan kan de oplossing u_0 ontwikkeld worden in een reeks die in H_K en in H convergeert, nl.

$$u_0 = \sum_{k=1}^{\infty} (f,K\phi_k)\phi_k.$$

bewijs:

zij $f \in H$ willekeurig; en zij u een element uit H_K. Uit de eigenschappen van H_K volgt dat (f,Ku) een begrensde lineaire functionaal is van u in H_K. De stelling van Riesz geeft: er is een en slechts een element $w \in H_K$, zodat voor alle $u \in H_K$ geldt:

(1.6) $$(f,Ku) = [w,u].$$

Hieruit volgt voor de functionaal F(u):

(1.7) $$F(u) = (Au,Ku) - (f,Ku) - (Ku,f) = [u,u] - [w,u] - [u,w] \text{ , voor } \forall u \in D(A).$$

Deze betrekking kunnen we blijkbaar gebruiken om F(u) te definiëren voor alle $u \in H_K$, en het minimum van F(u) te bepalen voor $u \in H_K$. Dit kunnen we eenvoudig uitvoeren.

Uit (1.7) volgt:

$F(u) = |u-w|^2 - |w|^2.$

Hieruit blijkt dat F(u) zijn minimum bereikt voor u = w, en dat

$\min_{u \in H_K} F(u) = -|w|^2.$

Verder geldt nu in H_K:

$w = \sum_{k=1}^{\infty} [w,\phi_k]\phi_k$ convergeert in H_K. Volgens (1.6) en ongelijkheid (1.5) volgt hieruit:

$w = \sum_{k=1}^{\infty} (f,K\phi_k)\phi_k$ convergeert in H_K en in H.

We zien dat het minimaliseren van F(u) een oplossing u_0 kan hebben met $u_0 \in H_K$, $u_0 \notin D(A)$, zodat het probleem Au = f geen oplossing heeft.

Daarom breiden we het definitiegebied D(A) van A (en daarmee A zelf) uit met de verzameling $\{u_0 | F(u)$ bereikt een minimum voor $u=u_0\}$. Dit voeren we als volgt uit: Door de betrekking (1.6) is een lineaire operator G gedefinieerd met D(G) = H en $R(G) \subset H_K$, die aan iedere $f \in H$ een element u_f toevoegt dat F(u) minimaliseert. We schrijven dan ook

(1.8) $\quad (f,Ku) = [Gf,u] \quad ,$ voor alle $f \in H$ en $u \in H_K$.

De operator G heeft de volgende eigenschappen:

1^0. G is begrensd als operator van H naar H_K en als operator van H naar H.

<u>bewijs</u>:

$|Gf|^2 = [Gf,Gf] = |(f,KGf)| \le ||f||.\,||KGf|| \le \frac{1}{\sqrt{\alpha}}\,||f||.|Gf|$ (zie (1.4)).

Dus $|Gf| \le \frac{1}{\sqrt{\alpha}}\,||f||$, en volgens (1.5): $||Gf|| \le \frac{1}{\alpha\beta}\,||f||$.

2^0. G heeft een inverse G^{-1}.

<u>bewijs</u>:

Stel Gf = 0. Dan volgt uit (1.8) dat (f,Ku) = 0 voor alle $u \in H_K$.
Uit de eigenschappen van K concluderen we dat $f \equiv 0$.
We kunnen nu de volgende stelling formuleren voor een K-p.d. operator A; voor K = I is deze stelling al geformuleerd door Friedrichs (zie o.a. [6]).

<u>stelling 1.4</u>

Als A een K-p.d. operator is, dan heeft A een gesloten K-p.d. uitbreiding, en wel de operator G^{-1}.

<u>bewijs</u> zie [9]; voor analoog bewijs als K = I, zie [6].

2. Gegeneraliseerde methode van Ritz

2.1. Formulering van de methode. Convergentie

Om de oplossing van het minimumprobleem van $F(u) = (Au,Ku) - (Ku,f) - (f,Ku)$ bij gegeven $f \in H$ te construeren, maken we gebruik van het volgende variatieprincipe:
laat $\{\phi_k\} \subset D(A)$ een rij lineair onafhankelijke elementen zijn die volledig is in H_K; zij verder

$w_n = \sum\limits_{j=1}^{n} \alpha_j^{(n)}\phi_j$ met $\alpha_j^{(n)}$ bepaald door de voorwaarde dat

$G(\alpha_1^{(n)},\dots,\alpha_n^{(n)}) \overset{\text{def}}{=} F(w_n)$ minimaal is in de deelruimte L opgespannen door $\phi_1, \dots, \phi_n$.
Deze voorwaarde leidt tot een algebraïsch stelsel, en wel als volgt:

$$F(w_n) = \sum_{j=1}^{n}\sum_{i=1}^{n} (A\phi_i,K\phi_j)\,\alpha_i^{(n)}\,\overline{\alpha_j^{(n)}} - \sum_{j=1}^{n}(K\phi_j,f)\,\alpha_j^{(n)} - \sum_{j=1}^{n}(f,K\phi_j)\,\overline{\alpha_j^{(n)}} .$$

$$G(\alpha_1^{(n)},\dots,\alpha_n^{(n)}) \text{ is minimaal} \Rightarrow \frac{\partial G}{\partial \alpha_j^{(n)}} = 0 \text{ voor } j = 1, 2, \dots, n. \Rightarrow$$

$$(2.1)\qquad \sum_{i=1}^{n} (A\phi_i,K\phi_j)\,\alpha_i^{(n)} = (f,K\phi_j) \qquad \text{voor } j = 1, 2, \dots, n.$$

Deze methode noemen we de gegeneraliseerde methode van Ritz.
We moeten nu laten zien, dat inderdaad geldt voor $n \to \infty$ $w_n \to w$ met $F(w)$ is minimaal. We geven het bewijs hiervoor.
Uit de eigenschappen van H_K volgt dat (f,Ku) voor iedere $f \in H$ een begrensde lineaire functionaal van u is in de Hilbertruimte H_K. Volgens de stelling van Riesz bestaat er een uniek element $w \in H_K$ zò dat geldt:

$$(2.2)\qquad (f,Ku) = [w,u] \quad \text{voor alle } u \in H_K ,$$

waarbij we eraan herinneren dat $[w,u] \overset{\text{def}}{=} (Aw,Ku)$.

Uit stelling 1.4 volgt dat $w \in D(A)$, met A voldoende uitgebreid, en tevens dat geldt $Aw = f$.

We beschouwen nu een stelsel lineair onafhankelijke elementen $\{\phi_j\} \subset D(A)$ dat volledig is in H_K.
Zij H_j de lineaire ruimte opgespannen door $\phi_1, \phi_2, \ldots, \phi_j$ en zij P_j de orthogonale projectie van H_K op H_j.
Zij verder:

(2.3) $$w_j = \sum_{l=1}^{j} \alpha_l^{(j)} \phi_l ,$$

waarin de coëfficiënten $\alpha_l^{(j)} (l = 1,2,\ldots,j)$ bepaald worden uit het feit dat:

(2.4) $$w_j = P_j w.$$

Nu geldt echter $w - P_j w \perp \phi_l$ voor $l = 1, 2, \ldots, j$ in de H_K-metriek,

d.w.z. $0 = [w-P_j w,\phi_l] = [w,\phi_l] - [P_j w,\phi_l] \qquad (l=1,2,\ldots,j) \implies$

(2.5) $$\sum_{i=1}^{j} (A\phi_i, K\phi_l)\, \alpha_i^{(j)} = (f, K\phi_l) \qquad (l=1,2,\ldots,j).$$

We zien dat de uitdrukkingen (2.1) en (2.5) identiek zijn.

Het stelsel $\{\phi_l\}_{l=1}^{j}$ is lineair onafhankelijk $\implies$ het stelsel $\{\alpha_l^{(j)}\}_{l=1}^{j}$ (en daarmee ook w_j) is eenduidig bepaald voor iedere j en gegeven f.
Uit de constructie van w_j volgt onmiddellijk $\lim_{j\to\infty} |w_j - w| = 0$.
Wegens de ongelijkheid (1.5) geldt dan $||w_j - w|| \leq \frac{1}{\gamma_2} |w_j - w| \to 0$ voor $j \to \infty$. Hiermee is de convergentie bewezen.
Als we bovendien aannemen dat $D(K) = D(A)$, kunnen we nog meer aantonen: ongelijkheid (1.4) impliceert de convergentie van Kw_n; omdat K gesloten is, geldt $Kw_n \to Kw$; uit ongelijkheid (1.2) volgt dan:

$$||Aw_n - f|| = ||Aw_n - Aw|| \leq \theta\, ||Kw_n - Kw|| \to 0 \text{ voor } n \to \infty .$$

We vatten deze resultaten samen in de volgende stelling:

<u>stelling 2.1</u>

Als A een K-p.d.operator is, en $\{\phi_l\} \subset D(A)$ is volledig in H_K, dan geldt:

1°. w_n is eenduidig bepaald voor iedere n door de gegeneraliseerde methode van Ritz, d.w.z. de uitdrukkingen (2.2) - (2.5);

2^o. w_n convergeert in H_K en in H naar de oplossing van $Aw = f$, met $f \in H$;
3^o. als ook nog geldt $D(K) = D(A)$, dan convergeert Aw_n naar f in H.

We beschouwen de volgende twee speciale gevallen:

1^o. $K = A$.

De operator A is inverteerbaar; de methode valt samen met de methode van de kleinste kwadraten. Stelling 2.1 geeft de convergentie van de methode:

$w_n \to w$ en $Aw_n \to f$, voor $n \to \infty$.

2^o. $K = I$.

De methode valt nu samen met de gewone methode van Ritz voor zelfgeadjungeerde positief definiete operatoren A.

Stelling 2.1 geeft weer de convergentie. Als A niet begrensd is, geldt niet altijd $Aw_n \to f$ in H. We moeten dan een extra eis opleggen aan A, zie o.a. [4] en [8], nl.:
de vergelijking $Aw = f$ in H heeft een oplossing w_0, waarvoor geldt dat (Aw_0, w_0) eindig is.
In dat geval construeren we volgens de methode van Ritz een rij $\{w_n\}$ met de eigenschap dat (Aw_n, w_n) eindig is voor iedere n.

2.2. Toepassingen

voorbeeld 1

Als eerste voorbeeld van een toepassing kiezen we de differentiaalvergelijking

$$(2.6)\qquad \begin{cases} -\dfrac{d^2}{dx^2}u = \dfrac{1}{(1+x)^2} & \text{in } (0,1)\ , \\ u = 0 \text{ voor } x = 0 \text{ en } x = 1, & \text{en} \quad u \in C^2(0,1)\ . \end{cases}$$

Zij $H = L_2(0,1)$; dan ligt $D(A)$ dicht in H.
De analytische oplossing van (2.6) is bekend, en luidt:

$$(2.7)\qquad u(x) = \ln(1+x) - x\ln(2).$$

De operator $A = -\dfrac{d^2}{dx^2}$ is zelfgeadjungeerd en positief definiet voor dit probleem. Dit is gemakkelijk in te zien door partiële integratie.

$$(Au,v) = -\int_0^1 v(x)\,\frac{d^2}{dx^2}\,u\,dx = [-v(x)\frac{du}{dx}]_0^1 + \int_0^1 (\frac{du}{dx}\,\frac{dv}{dx})\,dx = (Av,u).$$

$$u^2(x) = \left\{\int_0^x \frac{du}{dt}dt\right\}^2 \le x\int_0^x (\frac{du}{dt})^2 dt \le x\int_0^1 (\frac{du}{dt})^2 \implies (u,u) \le \frac{1}{2}\,(Au,u).$$

We kiezen in de Hilbertruimte $H_1 = H \cap \{u \mid u=0 \text{ voor } x=0 \text{ en } x=1\}$ het volledige stelsel coördinaatfuncties: $\phi_k(x) = x^k(1-x),\ k \ge 1$.
Vervolgens berekenen we de inprodukten $(A\phi_k, K\phi_l)$ en $(f,K\phi_l)$ met $k,l \ge 1$ en $f = \frac{1}{(1+x)^2}$, voor $K = A$ en $K = I$. Er gelden de volgende relaties:

$$(2.8)\qquad (A\phi_k,\phi_l) = \frac{kl}{k+l-1} - \frac{k(l+1)+l(k+1)}{k+l} + \frac{(k+1)(l+1)}{k+l+1},\quad k,l \ge 1.$$

$$(2.9)\qquad (A\phi_k,A\phi_l) = \frac{kl(l-1)(k-1)}{k+l-3} - \frac{k(l+1)(k-1)+l(k+1)(l-1)}{k+l-2} + \frac{(k+1)(l+1)kl}{k+l-1},\quad k+l > 3;$$

$$= 4,\quad \text{voor } k+l = 2;$$

$$= 2,\quad \text{voor } k+l = 3.$$

$$(2.10)\qquad \begin{cases} (f,\phi_k) = \displaystyle\int_0^1 \frac{x^k dx}{(1+x)^2} - \int_0^1 \frac{x^{k+1}}{(1+x)^2}\,dx, & k \ge 1; \\[2ex] (f,A\phi_k) = -k(k-1)\displaystyle\int_0^1 \frac{x^{k-2}}{(1+x)^2}\,dx + (k+1)k\int_0^1 \frac{x^{k-1}}{(1+x)^2}\,dx, & k \ge 1. \end{cases}$$

Zij $I_k = \int_0^1 \frac{x^k dx}{(1+x)^2}$, dan leiden we de volgende recurrente betrekking af voor I_k:

$$
(2.11) \qquad \begin{cases} I_k = \dfrac{1}{k-1} - 2I_{k-1} - I_{k-2}\,, & k > 1; \\[2ex] I_1 = \ln(2) - \dfrac{1}{2}\,,\ I_0 = \dfrac{1}{2}\,. & \end{cases}
$$

Nu rest nog het stelsel $\sum_{k=1}^{n} (A\phi_k, K\phi_l)\, \alpha_k^{(n)} = (f, K\phi_l)$ voor $l = 1, 2, \ldots, n$ op te lossen. Dit hebben we numeriek uitgevoerd voor toenemende waarden van n. Eerst beschouwen we het geval $K = I$, d.w.z. de gewone methode van Ritz. De volgende tabel bevat de waarden van de coëfficiënten $\alpha_k^{(n)}$ bij toenemende n ($n \leq 10, k \geq 1$).

n	$\alpha_1^{(n)}$	$\alpha_2^{(n)}$	$\alpha_3^{(n)}$	$\alpha_4^{(n)}$	$\alpha_5^{(n)}$	$\alpha_6^{(n)}$	$\alpha_7^{(n)}$	$\alpha_8^{(n)}$
1	+.2483							
2	+.2929	-.109						
3	+.3040	-.165	+.056					
4	+.3063	-.185	+.102	-.0309				
5	+.30675	-.191	+.126	-.0662	+.0176			
6	+.30683	-.1926	+.135	-.0898	+.0436	-.0104		
7	+.306849	-.1930	+.1387	-.1017	+.0652	-.0291	+.0065	
8	+.306852	-.1931	+.1398	-.1130	+.0792	-.0486	+.0202	-.004
9	+.306853	-.19318	+.1407	-.1230	+.1003	-.0913	+.0332	-.012
10	+.306853	-.19321	+.1416	-.189	+.1547	-.257	-.0199	+.041

We zien dat dit rekenproces niet leidt tot stabiele waarden van de coëfficiënten $\alpha_k^{(n)}$, met uitzondering van de eerste twee. Nu was dit ook niet te verwachten, omdat het stelsel $\{\phi_k\}$ niet minimaal is, d.w.z. een willekeurig element ϕ_{k_0} kan uit $\{\phi_k\}$ weggelaten worden zonder de volledigheid aan te tasten. Voor de noodzakelijkheid van de minimaliteit van $\{\phi_k\}$ voor stabiliteit zie [8]; dat $\{\phi_k | \phi_k(x) = x^k(1-x)\}$ inderdaad niet minimaal is in H_1, berust op de stelling van Müntz, die zegt dat $\{x^{p_k}\}$ volledig is in $L_2(0,1)$, als voor alle k $p_k \neq 0$, $p_k \to \infty$ voor $k \to \infty$, en $\sum_{k=1}^{\infty} p_k^{-1}$ divergeert (zie [3]).

We geven nu de resultaten voor het geval K = A, d.w.z. de methode van de de kleinste kwadraten.

n	$\alpha_1^{(n)}$	$\alpha_2^{(n)}$	$\alpha_3^{(n)}$	$\alpha_4^{(n)}$	$\alpha_5^{(n)}$	$\alpha_6^{(n)}$	$\alpha_7^{(n)}$	$\alpha_8^{(n)}$
1	+.25							
2	+.3068528194	-.1137						
3	+.3068528194	-.172	+.0584					
4	+.3068528194	-.188	+.1064	-.0320				
5	+.3068528194	-.1920	+.1286	-.0686	+.0183			
6	+.3068528194	-.1928	+.1367	-.0919	+.0452	-.0108		
7	+.3068528194	-.19310	+.1392	-.1030	+.0669	-.0301	+.007	
8	+.3068528194	-.19314	+.1399	-.1075	+.0797	-.0489	+.020	-.004
9	+.3068528194	-.19314	+.1400	-.1088	+.0848	-.0599	+.033	-.012
10	+.3068528194	-.19314	+.1398	-.1067	+.0740	-.0287	-.0199	+.041

Ook hier zien we dat met uitzondering van $\alpha_1^{(n)}$ en $\alpha_2^{(n)}$ de coëfficiënten geen limietwaarde bereiken voor toenemende n. Verder kunnen we uit de tabel aflezen dat voor n = 10 iets vreemds gebeurt bij $\alpha_7^{(10)}$ en $\alpha_8^{(10)}$: zij wisselen van teken.

Hieruit concluderen we dat voor n = 10 waarschijnlijk instabiliteit zal optreden.

We vragen ons natuurlijk af, of ondanks het instabiele karakter van het rekenproces, toch de oplossing van het probleem benaderd wordt. Dit hebben we gedaan door de analytische oplossing $u(x) = \ln(1+x) - x\ln(2)$ te vergelijken met de berekende benadering $u_n(x)$ voor $n = 1, 2, \ldots, 10$ in de punten $x = jh$ met $h = 1/10$ en $j = 1, 2, \ldots, 9$. Bovendien hebben we in deze punten de relatieve fout berekend, en over het interval (0,1) de maximale relatieve fout daaruit afgeleid.

We geven een vergelijkende tabel, waarin uitgezet zijn de maximale relatieve fout en het aantal coördinaatfuncties ϕ_k van beide methodes.

n	maximale relatieve fout bij de methode van Ritz	maximale relatieve fout bij de methode der kl.kwadr.
1	$.19_{10}{-0}$	$.25_{10}{-0}$
2	$.28_{10}{-1}$	$.61_{10}{-1}$
3	$.31_{10}{-2}$	$.67_{10}{-2}$
4	$.29_{10}{-3}$	$.89_{10}{-3}$
5	$.42_{10}{-4}$	$.13_{10}{-3}$
6	$.84_{10}{-5}$	$.15_{10}{-4}$
7	$.17_{10}{-5}$	$.22_{10}{-5}$
8	$.30_{10}{-6}$	$.28_{10}{-6}$
9	$.46_{10}{-6}$	$.66_{10}{-7}$
10	$.70_{10}{-6}$	$.23_{10}{-6}$

We kunnen het volgende aflezen uit deze tabel:

1°. de methode van Ritz convergeert in het begin sneller dan de methode van de kleinste kwadraten;

2°. de methode van de kleinste kwadraten convergeert het beste; het is waarschijnlijk dat dit komt, omdat de methode minder instabiel is dan de methode van Ritz.

De grootste precisie wordt bereikt voor n = 9, met de methode der kleinste kwadraten, en wel $.66_{10}{-7}$.

Omdat we op het interval (0,1) werken, kunnen we ons afvragen, in hoeverre de verkregen benadering $u_9(x)$ samenhangt met de Taylorontwikkeling van de analytische oplossing u(x) in 0:

$$u(x) = \ln(1+x) - x\ln(2) = \sum_{m=1}^{\infty} (-1)^{m+1} \frac{x^m}{m} - x\ln(2) = \tag{2.12}$$

$$(1-\ln(2))x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \ldots$$

en

$$(2.13)\qquad u_9(x) = \alpha_1^{(9)}x(1-x) + \alpha_2^{(9)}x^2(1-x) + \ldots + \alpha_9^{(9)}x^9(1-x) =$$

$$\alpha_1^{(9)}x - (\alpha_2^{(9)}-\alpha_1^{(9)})x^2 + (\alpha_3^{(9)}-\alpha_2^{(9)})x^3 - \ldots = \sum_{k=1}^{9} b_k x^k.$$

Het blijkt dat voor k = 5, de coëfficiënten b_k die van de Taylorontwikkeling van u(x) in 0 goed benaderen; voor de eerste coëfficiënt geldt zelfs een relatieve precisie van 10^{-9}.

We geven nu een overzicht der b_k's in de volgende tabel.

k	b_k	Taylorreeks	abs.fout	rel.fout
1	.3068528194	.3068528195	10^{-10}	$.3\,10^{-9}$
2	.499996	.5	$.4\,10^{-5}$	$.8\,10^{-5}$
3	.3332	.3333333333	10^{-4}	$.3\,10^{-3}$
4	.2489	.25	$.11\,10^{-2}$	$.4\,10^{-2}$
5	.1936	.2	$.64\,10^{-2}$	$.3\,10^{-1}$

voorbeeld 2

Als tweede voorbeeld beschouwen we het probleem

$$(2.14)\qquad \begin{cases} \dfrac{d^3u}{dx^3} = \dfrac{2}{(1+x)^3}, & \text{in } (0,1), \\ u = 0 \text{ voor } x = 0, & \\ u' = 0 \text{ voor } x = 0 \text{ en } x = 1, & u \in C^3(0,1). \end{cases}$$

Zij $H = L_2(0,1)$, dan weten we dat $D(A) = C^3(0,1)$ dicht ligt in H.
De analytische oplossing van (2.14) luidt:

$$(2.15)\qquad u(x) = \ln(1+x) + \tfrac{1}{4}x^2 - x.$$

De operator $A = -\dfrac{d^3}{dx^3}$ is in dit probleem niet meer zelfgeadjungeerd, zodat de gewone methode van Ritz (K=I) niet toegepast kan worden om de oplossing van (2.14) te benaderen.

In het eerste hoofdstuk hebben we gezien dat er een operator K bestaat, nl. $K = \dfrac{d}{dx}$, waarvoor A K-p.d. is. We kunnen dus de gegeneraliseerde methode voor niet-triviale K toepassen. Verder zullen we de resultaten vergelijken met die van de methode van de kleinste kwadraten (K = A).

We moeten nog wel laten zien dat het inprodukt (Au,Kv) symmetrisch is voor alle u,v in de ruimte $H_1 = L_2(0,1) \cap \{u \mid u(0) = 0 \text{ en } u'(0) = u'(1) = 0\}$.

Voor K = A is dit geen probleem; voor $K = \dfrac{d}{dx}$ geldt:

$$(2.16) \qquad (Au,Kv) = \int_0^1 -\frac{d^3u}{dx^3}\frac{dv}{dx}\,dx = \left[-\frac{d^2u}{dx^2}\frac{dv}{dx}\right]_0^1 + \int_0^1 \frac{d^2u}{dx^2}\frac{d^2v}{dx^2}\,dx =$$

$$(Ku,Av).$$

In H_1 kiezen we weer een volledig stelsel $\{\phi_k\}$, nl. het stelsel $\phi_k(x) = x^{k+1}\left(\frac{1}{k+1} - \frac{x}{k+2}\right)$, $k \geq 1$. Dat dit stelsel volledig is, berust op het feit dat voor $w = \dfrac{du}{dx}$ het probleem (2.6) terugkeert met randvoorwaarden w = 0 voor x = 0 en x = 1.

Op analoge wijze als in het eerste voorbeeld rekenen we nu de inprodukten uit. Voor $(A\phi_k,K\phi_1)$ zie relatie (2.8), als $K = \dfrac{d}{dx}$, en (2.9), als K = A. Voor $(f,K\phi_k)$ met $f = \dfrac{-2}{(1+x)^3}$ geldt:

$$(2.17) \quad \begin{cases} \left(f,\dfrac{d\phi_k}{dx}\right) = -2\displaystyle\int_0^1 \frac{x^k}{(1+x)^3}\,dx + 2\int_0^1 \frac{x^{k+1}}{(1+x)^3}\,dx, & k \geq 1; \\ \\ (f,A\phi_k) = 2k(k-1)\displaystyle\int_0^1 \frac{x^{k-2}}{(1+x)^3}\,dx - 2(k+1)k\int_0^1 \frac{x^{k-1}}{(1+x)^3}\,dx, & k \geq 1. \end{cases}$$

Zij $I_k = \displaystyle\int_0^1 \frac{x^k}{(1+x)^3}\,dx$, dan geldt de betrekking:

$$(2.18)\qquad \begin{cases} I_k = \frac{1}{k-2} - 3I_{k-1} - 3I_{k-2} - I_{k-3}\,, & k > 2; \\ I_0 = 3/8;\ I_1 = 1/8;\ I_2 = \ln(2) - 5/8\,. \end{cases}$$

Ten slotte lossen we het stelsel vergelijkingen

$\sum_{k=1}^{n} (A\phi_k, K\phi_l)\, \alpha_k^{(n)} = (f, K\phi_l)$ op voor $l = 1, 2, \ldots, n$, met numerieke procedures.

Eerst geven we de resultaten voor het geval $K = \frac{d}{dx}$, de gegeneraliseerde methode van Ritz.

n	$\alpha_1^{(n)}$	$\alpha_2^{(n)}$	$\alpha_3^{(n)}$	$\alpha_4^{(n)}$	$\alpha_5^{(n)}$	$\alpha_6^{(n)}$	$\alpha_7^{(n)}$	$\alpha_8^{(n)}$
1	-.3411							
2	-.45787	+.2335						
3	-.48987	+.3935	-.1600					
4	-.49771	+.4641	-.3246	+.1097				
5	-.499505	+.48916	-.4250	+.2602	-.0753			
6	-.499896	+.49698	-.4719	+.3775	-.2043	+.0516		
7	-.499978	+.499185	-.4903	+.4450	-.3258	+.1569	-.0351	
8	-.499990	+.499623	-.4951	+.4692	-.3886	+.2448	-.0979	+.0180
9	-.499978	+.498835	-.4839	+.3958	-.1298	-.2761	+.5006	-.3471

Net als in het eerste voorbeeld zien we instabiliteit optreden (voor $n = 9$). De resultaten van de methode van de kleinste kwadraten, $K = A$, volgen nu in tabelvorm.

n	$\alpha_1^{(n)}$	$\alpha_2^{(n)}$	$\alpha_3^{(n)}$	$\alpha_4^{(n)}$	$\alpha_5^{(n)}$	$\alpha_6^{(n)}$	$\alpha_7^{(n)}$	$\alpha_8^{(n)}$
1	-.3750							
2	-.5000	+.2500						
3	-.499999999995	+.4194	-.1694					
4	-.499999999986	+.4771	-.3426	+.1154				
5	-.499999999977	+.4940	-.4384	+.2732	- .0789			
6	-.499999999968	+.4985	-.4788	+.3901	- .2138	+ .0540		
7	-.499999999959	+.49966	-.4934	+.4539	- .3381	+ .1648	- .0370	
8	-.499999999955	+.49992	-.4981	+.4828	- .4211	+ .2869	- .1258	+ .0254
9	-.499999999950	+.499991	-.49967	+.4956	- .4714	+ .3944	- .2528	+ .1035
10	-.499999999923	+.500057	-.50158	+.5155	- .5741	+ .6911	- .7565	+ .6020
11	-.499999999495	+.500181	-.50660	+.5858	-1.057	+2.5706	-5.167	+6.972

Ook nu zien we weer hetzelfde beeld als in het vorige voorbeeld:voor toenemende n bereiken alleen de eerste coëfficiënten een limietwaarde, de andere niet. De instabiliteit is duidelijk te zien voor n = 11 met K = A en voor n = 9 met $K = \frac{d}{dx}$.

Net als in het eerste voorbeeld gaan we na, of de verkregen benaderingen de oplossing van (2.14) goed representeren. Dit voeren we op analoge wijze uit, door in de punten x = jh, j = 1, 2, ..., 9 en h = 1/10 de benaderingen te vergelijken met de analytische oplossing. Hieruit berekenen we vervolgens de maximale relatieve fout. De beide methodes worden met elkaar vergeleken in de volgende tabel.

n	max.rel.fout voor $K = \frac{d}{dx}$	max.rel.fout voor K = A
1	$.27_{10}{-0}$	$.20_{10}{-0}$
2	$.57_{10}{-1}$	$.30_{10}{-1}$
3	$.97_{10}{-2}$	$.12_{10}{-1}$
4	$.14_{10}{-2}$	$.20_{10}{-2}$
5	$.14_{10}{-3}$	$.36_{10}{-3}$
6	$.11_{10}{-4}$	$.58_{10}{-4}$
7	$.13_{10}{-5}$	$.79_{10}{-5}$
8	$.63_{10}{-6}$	$.79_{10}{-6}$
9	$.29_{10}{-5}$	$.92_{10}{-7}$
10	$.81_{10}{-4}$	$.30_{10}{-6}$

We zien hetzelfde als in het eerste voorbeeld; voor n = 9 wordt de grootste precisie bereikt, n.l. $.92_{10}{-7}$ door de methode van de kleinste kwadraten.

Als laatste vergelijken we deze benadering met de Taylorreeks van de analytische oplossing u(x) van (2.14) in het punt x = 0:

$$u(x) = \ln(1+x) + \frac{1}{4}x^2 - x = \sum_{k=1}^{\infty} (-1)^{k+1} \frac{x^k}{k} + \frac{1}{4}x^2 - x = -\frac{1}{4}x^2 + \frac{1}{3}x^3 - \frac{1}{4}x^4 + \ldots, \tag{2.19}$$

$$u_9(x) = \sum_{k=1}^{9} \alpha_k x^{k+1} \left(\frac{1}{k+1} - \frac{x}{k+2}\right) = \sum_{k=1}^{9} b_k \, x^{k+1} .$$

k	b_k	Taylorreeks	abs.fout	rel.fout
1	-.2499999999975	-.25	$.25_{10}{-11}$	$_{10}{-11}$
2	+.3333303	+.33333333	$.30_{10}{-5}$	$.90_{10}{-5}$
3	-.249915	-.25	$.85_{10}{-4}$	$.34_{10}{-3}$
4	+.19905	+.2	$.95_{10}{-3}$	$.47_{10}{-2}$
5	-.1628	-.1666666	$.38_{10}{-2}$	$.23_{10}{-1}$

3. Bepaling van eigenwaarden

3.1. Inleiding. Stellingen

De eigenwaarden van een operator spelen een belangrijke rol in de toegepaste wiskunde. Vele problemen uit de mechanica geven aanleiding tot differentiaalvergelijkingen, waarvan de oplossing gegeven wordt door de z.g. eigentrillingen en eigenfrequenties (d.w.z. eigenfuncties en eigenwaarden van de betreffende differentiaaloperator). Ook bij het numerieke oplossen van differentiaalvergelijkingen is met name de kleinste eigenwaarde van belang.

In dit hoofdstuk zullen we laten zien dat het probleem van het bepalen van de eigenwaarden van een K-p.d.operator terug te voeren is tot een minimaliseringsprobleem; hiervan kan de oplossing bepaald worden met de gegeneraliseerde methode van Ritz. De stellingen vormen een generalisatie van de desbetreffende stellingen voor zelfgeadjungeerde operatoren (zie o.a. Mikhlin [6], [7]).

Anders dan in hoofdstuk 1 en 2 beschouwen we uitsluitend een reële Hilbertruimte H die separabel is, en een operator A gedefinieerd op een dicht gebied $D(A) \subset H$.

Zij verder K net als in hoofdstuk 1 een gesloten operator, gedefinieerd op een dicht gebied $D(K) \supseteq D(A)$ in H met de in hoofdstuk 1 vermelde eigenschappen. Laat nu verder nog gelden:

1°. K is zelfgeadjungeerd;

2°. A is K-positief definiet, d.w.z. er bestaat een $\alpha > 0$, zodat voor alle $u \in D(A)$ geldt $(Au,Ku) \geq \alpha||Ku||^2$;

3°. voor alle $u,v \in D(A)$ $(Au,Kv) = (Ku,Av)$ (symmetrie).

$D(A)$ kan uitgebreid worden tot een Hilbertruimte H_K. A kan uitgebreid worden tot een gesloten operator. Laten we veronderstellen dat dit gebeurd is.

We kunnen nu de volgende eigenschappen voor eigenwaarden afleiden:

(i). De eigenwaarden van een K-p.d.operator A met K zelfgeadjungeerd, zijn reëel.

<u>bewijs</u>:

zij u eigenvector van A met eigenwaarde λ.

Dan is $Au = \lambda u \Rightarrow (Au,Ku) = (\lambda u,Ku) = \lambda(u,Ku) \Rightarrow \lambda = \dfrac{(Au,Ku)}{(u.Ku)}$.

Nu is (Au,Ku) reëel, omdat A K-p.d. is, maar ook (u,Ku) is reëel, omdat K zelfgeadjungeerd is.

(ii). De eigenwaarden van een positief definiete operator zijn positief;

(iii). Voor eigenfuncties u,v van een K-p.d.operator A met K zelfgeadjungeerd, die tot verschillende eigenwaarden behoren, geldt dat $(u,Kv) = 0$.

bewijs:

stel $Au = \lambda u$ en $Av = \mu v$; wegens (i) zijn λ en μ reëel; verder is $\lambda \neq \mu$. Nu volgt: $(Au,Kv) = \lambda(u,Kv)$ en

$(Ku,Av) = \mu(Ku,v)$.

Wegens de symmetrie van (Au,Kv) en de zelfgeadjungeerdheid van K volgt hieruit:

$\lambda(u,Kv)=(Au,Kv)=(Ku,Av)=\mu(Ku,v)=\mu(u,Kv) \Longrightarrow (\lambda-\mu)(u,Kv) = 0.$

Bij het bewijs van (i) hebben we gezien dat de eigenwaarde λ van A te vinden is onder de waarden van de functionaal

$$(3.1) \qquad G(u) \overset{\text{def}}{=} \frac{(Au,Ku)}{(u,Ku)}, \quad \text{voor alle } u \in H_K .$$

Omdat α, $\beta > 0$ bestaan met $(Au,Ku) \geq \alpha||Ku||^2$ en $||Ku|| \geq \beta||u||$ voor alle $u \in D(A)$ en dit uit te breiden is voor alle $u \in H_K$ (zie (1.3) en (1.4)), is $G(u)$ naar beneden begrensd in H_K. Dus $G(u)$ bezit een minimum m in H_K.

stelling 3.1

Laat A een K-p.d.operator zijn en K zelfgeadjungeerd. Zij verder $\min_{u \in H_K} G(u) = m$.

Dan geldt: als er een $u_0 \neq 0$ in H_K bestaat met $m = G(u_0)$, dan is m de kleinste eigenwaarde van de operator A en u_0 de bijbehorende eigenfunctie.

bewijs:

zij $\eta \in H_K$ willekeurig, en t een reëel getal. Beschouw $\phi = \phi(t)$ met

$$\phi(t) = G(u_0+t\eta) = \frac{t^2(A\eta,K\eta) + 2t(Au_0,K\eta) + (Au_0,Ku_0)}{t^2(\eta,K\eta) + 2t(u_0,K\eta) + (u_0,Ku_0)} .$$

Omdat $G(u)$ voor $u = u_0$ minimaal is, geldt dat $\phi'(0) = 0$. Hieruit volgt na enig rekenwerk:

$$(3.2) \qquad (u_0,Ku_0)(Au_0,K\eta) - (Au_0,Ku_0)(u_0,K\eta) = 0.$$

Nu is $G(u_0) = (Au_0,Ku_0)/(u_0,Ku_0) = m$. Substitutie hiervan in (3.2) levert:

$$(3.3) \qquad (Au_0,K\eta) - (mu_0,K\eta) = (Au_0-mu_0,K\eta) = 0 \ , \quad \text{voor alle } \eta \in H_K.$$

Omdat voor alle $u \in H_K$ $||Ku|| \geq \beta||u||$ en $R(K)$ dicht in H ligt, concluderen we dat $Au_0 = mu_0$.

Volgens deze stelling kunnen we de kleinste eigenwaarde bepalen door het minimum te zoeken van $G(u)$, zie (3.1). Voor de andere eigenwaarden leiden we de volgende stelling af:

stelling 3.2

Laten $\lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_n$ n op elkaar volgende eigenwaarden zijn van de K-p.d.operator A waar K zelfgeadjungeerd is, en $u_1, u_2, \ldots, u_n$ de bijbehorende eigenvectoren, die zo georthonormeerd zijn dat $(u_l,Ku_m) = 0$ voor $l \neq m$, $l,m = 1, 2, \ldots, n$ en $(u_l,Ku_l) = 1$ voor $1 \leq l \leq n$.
Laat verder een vector $u_{n+1} \neq 0$ bestaan die (3.1) minimaliseert onder de nevenvoorwaarden $(u_{n+1},Ku_l) = 0$ voor $1 \leq l \leq n$. Dan is u_{n+1} een eigenvector van de operator A die bij de eigenwaarde λ_{n+1} behoort die de eerstvolgende in grootte is na λ_n, met de eigenschap dat:

$$(3.4) \qquad \lambda_{n+1} = \frac{(Au_{n+1},Ku_{n+1})}{(u_{n+1},Ku_{n+1})} \ .$$

bewijs:

zij η een willekeurig element uit H_K; dan beschouwen we het verschil $v = \eta - \sum_{j=1}^{n} (\eta,Ku_j)u_j$. Voor $1 \leq l \leq n$ geldt wegens de eigenschappen van $\{u_k\}_{k=1}^{n}$:

$$(v,Ku_l) = (\eta,Ku_l) - \sum_{j=1}^{n} (\eta,Ku_j)(u_j,Ku_l) = (\eta,Ku_l) - (\eta,Ku_l) = 0.$$

Zij vervolgens t een reëel getal. Beschouw $u_{n+1} + tv$, waarin u_{n+1} aan de voorwaarden van de stelling voldoet. Uit dezelfde overwegingen als in stelling 3.1 leiden we af dat:

(3.5) $\quad (Au_{n+1}-\lambda_{n+1}u_{n+1},Kv) = 0$, zie ook (3.3), voor alle $\eta \in H_K$.

Hieruit volgt:

$$(Au_{n+1}-\lambda_{n+1}u_{n+1},K\eta) = (Au_{n+1}-\lambda_{n+1}u_{n+1},Kv) +$$

$$\sum_{j=1}^{n} (\eta,Ku_j)\ (Au_{n+1}-\lambda_{n+1}u_{n+1},Ku_j).$$

Substitutie van (3.5) in deze betrekking geeft:

$$(3.6) \qquad (Au_{n+1}-\lambda_{n+1}u_{n+1},K\eta) = \sum_{j=1}^{n} (\eta,Ku_j)(Au_{n+1}-\lambda_{n+1}u_{n+1},Ku_j).$$

Verder weten we dat voor $1 \le j \le n$ geldt:

$$(Au_{n+1}-\lambda_{n+1}u_{n+1},Ku_j) = (Au_{n+1},Ku_j) - \lambda_{n+1}(u_{n+1},Ku_j) = (Au_{n+1},Ku_j) .$$

Omdat K zelfgeadjungeerd is en (Au,Kv) symmetrisch is, volgt:

$$(Au_{n+1},Ku_j) = (Ku_{n+1},Au_j) = \lambda_j(Ku_{n+1},u_j) = \lambda_j(u_{n+1},Ku_j) = 0, \quad 1 \le j \le n.$$

Dit gesubstitueerd in (3.6) levert ten slotte:

$$(3.7) \qquad (Au_{n+1}-\lambda_{n+1}u_{n+1},K\eta) = 0 \quad \text{voor alle } \eta \in H_K .$$

Wegens de eigenschappen van K volgt hieruit dat $Au_{n+1} = \lambda_{n+1}u_{n+1}$.
We moeten nu nog bewijzen dat λ_{n+1} inderdaad de eerstvolgende eigenwaarde is na λ_n, d.w.z. $\lambda_{n+1} \ge \lambda > \lambda_n \Rightarrow \lambda = \lambda_{n+1}$.

Stel v is de bijbehorende eigenvector van λ; dan geldt $(v,Ku_k) = 0$ voor $1 \le k \le n$. Dus:

$$\lambda = \frac{(Av,Kv)}{(v,Kv)} \ge \min_{u\in H_K} \frac{(Au,Ku)}{(u,Ku)} = \lambda_{n+1}.$$

We zijn er in de voorgaande stellingen van uitgegaan dat de operator A een discreet spectrum van eigenwaarden bezit. De existentie hiervan is een apart probleem, dat we niet zullen behandelen. Wel vermelden we een stelling hierover die op analoge wijze als in [6], stelling 2, pag.34, bewezen kan worden:

stelling 3.3

Als A K-p.d. is en K is zelfgeadjungeerd, en als iedere begrensde verzameling in H_K compact is in H, dan geldt:

1°. A heeft een aftelbare verzameling L van eigenwaarden, mits H oneindig dimensionaal is;

2°. het enige verdichtingspunt van L is oneindig;

3°. de rij der eigenvectoren van A is volledig in H_K en in H.

In de volgende paragraaf zullen we de eigenwaarden van een K-p.d.operator A bepalen door het minimaliseren van (3.1).

3.2. Gegeneraliseerde methode van Ritz, toegepast op het eigenwaardeprobleem

We zoeken het minimum van $G(u) = \frac{(Au,Ku)}{(u,Ku)}$ voor alle $u \in H_K$. Daartoe kiezen we in H_K een volledig stelsel $\{\phi_l\}$.

Zij $u_n = \sum_{j=1}^{n} \alpha_j^{(n)} \phi_j$, waar de $\alpha_j^{(n)}$'s($j = 1,2,\dots,n$) bepaald worden door de voorwaarde:

$G(u_n)$ is minimaal in de deelruimte L van H_K opgespannen door $\{\phi_l\}_{l=1}^{n}$.

We definiëren nu de functionaal $H(\alpha_1^{(n)},\dots,\alpha_n^{(n)}) \overset{\text{def}}{\equiv} G(u_n)$, voor vaste n. Een noodzakelijke voorwaarde opdat G minimaal is, is dat de partiële afgeleiden van G als functie van $\alpha_1^{(n)},\dots,\alpha_n^{(n)}$, d.w.z. de afgeleiden van H alle 0 zijn:

$$\frac{\partial H}{\partial \alpha_k^{(n)}} = 0 \text{ voor } k = 1,2,\dots,n.$$

We werken deze voorwaarden nader uit. Er geldt:

$$H(\alpha_1^{(n)},\dots,\alpha_n^{(n)}) = \sum_{j,k=1}^{n} (A\phi_j,K\phi_k)\alpha_j^{(n)}\alpha_k^{(n)} \Big/ \sum_{j,k=1}^{n} (\phi_j,K\phi_k)\alpha_j^{(n)}\alpha_k^{(n)}. \tag{3.8}$$

Het minimum van (3.8) bepalen is equivalent met het minimaliseren van:

$$\sum_{j,k=1}^{n} (A\phi_j,K\phi_k)\alpha_j^{(n)}\alpha_k^{(n)} - \lambda \sum_{j,k=1}^{n} (\phi_j,K\phi_k)\alpha_j^{(n)}\alpha_k^{(n)}, \tag{3.9}$$

waarin λ de multiplicator van Lagrange voorstelt.

Differentiatie van (3.9) naar $\alpha_k^{(n)}$ ($1 \leq k \leq n$) leidt tot het volgende stelsel vergelijkingen:

$$\sum_{j=1}^{n} \Big[(A\phi_j,K\phi_k) - \lambda(\phi_j,K\phi_k) \Big] \alpha_j^{(n)} = 0 \text{ voor } k = 1,2,\dots,n; \tag{3.10a}$$

of in vectornotatie:

$$(B - \lambda C)\,\vec{\alpha} = 0, \tag{3.10b}$$

zodat $C^{-1}B\vec{\alpha} = \lambda\vec{\alpha}$.

Hierin is $B_{ij} = (A\phi_i,K\phi_j)$ en $C_{ij} = (\phi_i,K\phi_j)$ voor $j = 1,2,\dots,n$,

en $1 \leq i \leq n$.
Hieruit blijkt dat de multiplicator λ een eigenwaarde is van $C^{-1}B$. We bewijzen nu dat de functionaal $G(u)$ minimum λ heeft in L:
Laat λ voldoen aan (3.10a) en α_j $(1 \leq j \leq n)$ de oplossing zijn van (3.10a). Vermenigvuldiging van (3.10a) met α_k en sommatie over $k = 1,2,\ldots,n$ leidt tot de volgende uitdrukking:

$$\sum_{j,k=1}^{n} \left\{ (A\phi_j, K\phi_k) - \lambda(\phi_j, K\phi_k) \right\} \alpha_j \alpha_k = 0 \Rightarrow$$

$$\lambda = \sum_{j,k=1}^{n} (A\phi_j, K\phi_k)\alpha_j\alpha_k \Big/ \sum_{j,k=1}^{n} (\phi_j, K\phi_k)\alpha_j\alpha_k = \frac{(Au_n, Ku_n)}{(u_n, Ku_n)},$$

met $u_n = \sum_{j=1}^{n} \alpha_j \phi_j$.

We zien dat de gegeneraliseerde methode van Ritz bovengrenzen geeft voor de eigenwaarden van A. Er kan bewezen worden dat voor toenemende n deze bovengrenzen inderdaad naar de werkelijke waarden van de eigenwaarden convergeren (zie o.a. [7]).

3.3. Toepassingen

voorbeeld 1

Zij $H = L_2(0,1)$, en zij $A = -\dfrac{d^2}{dx^2}$, zodat $D(A) = C^2(0,1)$ dicht in H ligt. We nemen $K = I$ en ook $K = A$, zodat aan de voorwaarden die aan het begin van dit hoofdstuk staan, voldaan is. We stellen het probleem:

$$(3.11) \quad \begin{cases} Au = \lambda u \text{ in } (0,1), \\ \\ u = 0 \text{ voor } x = 0 \text{ en } x = 1. \end{cases}$$

Dan is A symmetrisch en positief definiet. De eigenfuncties en eigenwaarden van A zijn bekend voor dit probleem:

$$\begin{cases} u_n = \sin(n\pi x), \\ \lambda_n = n^2\pi^2, \ (n = 1,2,\ldots). \end{cases}$$

Het stelsel $\{x^k(1-x) \mid k \geq 1\}$ is volledig in de ruimte $H_1 = H \cap \{u \mid u(0) = u(1) = 0\}$. Voor de berekening van de eigenwaarden is het nodig de inprodukten $(A\phi_k, K\phi_l)$ en $(\phi_k, K\phi_l)$ te berekenen voor $k, l = 1,2,\ldots$ en $K = I$, $K = A$.
Gedeeltelijk is dit al gebeurd in het vorige hoofdstuk (zie (2.8) en (2.9)). Verder geldt nog:

$$(3.12) \qquad (\phi_k, \phi_l) = \frac{1}{k+l+1} - \frac{2}{k+l+2} + \frac{1}{k+l+3} \text{ met } k,\ l \geq 1.$$

We beschouwen het volgende stelsel vergelijkingen:

$$(3.13) \qquad \sum_{j=1}^{n} \left[(A\phi_j, K\phi_l) - \lambda(\phi_j, K\phi_l) \right] \alpha_j^{(n)} = 0 \text{ voor } l = 1,2,\ldots,n,$$

of in vectornotatie: $C^{-1}B\vec{\alpha} = \lambda\vec{\alpha}$, zie ook (3.10a) en (3.10b).

Het bepalen van de eigenwaarden van $C^{-1}B$ is met behulp van numerieke procedures uitgevoerd voor toenemend aantal coördinaatfuncties ϕ_k.
Het blijkt dat het conditiegetal van de matrices $C^{-1}B$ groeit voor toenemende dimensie, zodat er snel instabiliteit optreedt in het rekenproces. We geven eerst de resultaten voor $K = I$, de gewone methode van Ritz.

n	$\lambda_1^{(n)}$	$\lambda_2^{(n)}$	$\lambda_3^{(n)}$	$\lambda_4^{(n)}$	$\lambda_5^{(n)}$	$\lambda_6^{(n)}$	$\lambda_7^{(n)}$
1	10						
2	10	42					
3	9.8697	42	102.13				
4	9.8697	39.50	102.13	200.49			
5	9.869604	39.50	89.17	200.49	350.95		
6	9.869599	39.478	89.17	159.99	350.95	570.59	
7	9.869472	39.446	88.94	160.07	253.87	570.52	877.59
$\lambda_k = k^2\pi^2$	9.8696044	39.4784	88.826	157.91	246.74	355.31	483.61

Vervolgens berekenen we de precisie waarmee λ_k $(k = 1,2,\ldots)$ benaderd wordt.

n	$\lambda_1^{(n)}$	$\lambda_2^{(n)}$	$\lambda_3^{(n)}$	$\lambda_4^{(n)}$	$\lambda_5^{(n)}$	$\lambda_6^{(n)}$	$\lambda_7^{(n)}$
1	$.13_{10}^{-1}$						
2	$.13_{10}^{-1}$	$.64_{10}^{-1}$					
3	$.15_{10}^{-4}$	$.64_{10}^{-1}$	$.15_{10}^{-0}$				
4	$.15_{10}^{-4}$	$.59_{10}^{-3}$	$.15_{10}^{-0}$	$.27_{10}^{-0}$			
5	$.84_{10}^{-7}$	$.59_{10}^{-3}$	$.39_{10}^{-2}$	$.27_{10}^{-0}$	$.42_{10}^{-0}$		
6	$.50_{10}^{-6}$	$.41_{10}^{-5}$	$.39_{10}^{-2}$	$.13_{10}^{-1}$	$.42_{10}^{-0}$	$.61_{10}^{-0}$	
7	$.13_{10}^{-4}$	$.24_{10}^{-3}$	$.13_{10}^{-2}$	$.13_{10}^{-1}$	$.29_{10}^{-1}$	$.61_{10}^{-0}$	$.81_{10}^{-0}$

Nu dan de resultaten voor het geval K = A, de methode van de kleinste kwadraten.

n	$\lambda_1^{(n)}$	$\lambda_2^{(n)}$	$\lambda_3^{(n)}$	$\lambda_4^{(n)}$	$\lambda_5^{(n)}$	$\lambda_6^{(n)}$	$\lambda_7^{(n)}$
1	12						
2	12	60					
3	9.87509	60	170.1				
4	9.87509	39.76	170.1	380.2			
5	9.869606	39.76	91.2	380.2	738.9		
6	9.869606	39.479	91.2	167.9	738.9	1304.6	
7	9.86962	39.479	88.86	167.9	275.7	1304.6	2145.8
8	9.86898	39.455	89.07	158.8	274.5	421.1	2145.8
$\lambda_k = k^2\pi^2$	9.869604	39.4784	88.826	157.91	246.74	355.31	483.6

Vervolgens berekenen we, net als in het geval K = I, uit deze gegevens de relatieve fout.

n	$\lambda_1^{(n)}$	$\lambda_2^{(n)}$	$\lambda_3^{(n)}$	$\lambda_4^{(n)}$	$\lambda_5^{(n)}$	$\lambda_6^{(n)}$	$\lambda_7^{(n)}$
1	$.22_{10}{-0}$						
2	$.22_{10}{-0}$	$.52_{10}{-0}$					
3	$.56_{10}{-3}$	$.52_{10}{-0}$	$.92_{10}{-0}$				
4	$.56_{10}{-3}$	$.73_{10}{-2}$	$.92_{10}{-0}$	$.14_{10}{+1}$			
5	$.25_{10}{-6}$	$.73_{10}{-2}$	$.27_{10}{-1}$	$.14_{10}{+1}$	$.20_{10}{+1}$		
6	$.17_{10}{-6}$	$.32_{10}{-4}$	$.27_{10}{-1}$	$.63_{10}{-1}$	$.20_{10}{+1}$	$.27_{10}{+1}$	
7	$.19_{10}{-5}$	$.30_{10}{-4}$	$.39_{10}{-3}$	$.63_{10}{-1}$	$.12_{10}{-0}$	$.27_{10}{+1}$	$.34_{10}{+1}$
8	$.64_{10}{-4}$	$.61_{10}{-3}$	$.27_{10}{-2}$	$.55_{10}{-2}$	$.11_{10}{-0}$	$.19_{10}{-0}$	$.34_{10}{+1}$

Als we deze resultaten vergelijken met die van de gewone methode van Ritz, dan blijkt dat de methode van Ritz de hoogste precisie haalt; beide methodes zijn in dit voorbeeld even gevoelig voor de stabiliteit van het rekenproces. Het blijkt dat de instabiliteit af te lezen ~~valt~~ valt van het feit dat de benadering slechter wordt in plaats van beter bij toenemende n; in dit geval voor n = 7.

<u>voorbeeld 2</u>

Zij $H = L_2(0,1)$ en $A = \frac{d^4}{dx^4}$, zodat $D(A) = C^4(0,1)$ dicht ligt in H. We beschouwen het probleem:

$$(3.14)\qquad \begin{cases} \dfrac{d^4u}{dx^4} = \lambda u \text{ in } (0,1), \\ u(0) = u'(0) = 0, \\ u''(1) = u'''(1) = 0, \text{ met } u \in C^4(0,1). \end{cases}$$

Door partiële integratie kunnen we afleiden dat de operator A symmetrisch is en positief definiet. We zullen alleen de methode van Ritz (K = I) toepassen, omdat de methode van de kleinste kwadraten vergeleken met de methode van Ritz veel meer rekenwerk met zich meebrengt.
Er geldt:

$$(3.15)\qquad (Au,v) = \int_0^1 \left(\frac{d^2u}{dx^2}\,\frac{d^2v}{dx^2}\right)dx.$$

De eigenwaarden λ van A voldoen aan de volgende transcendente vergelijking:

(3.16) $$\cos(\nu)\cosh(\nu) = -1, \text{ met } \nu = \lambda^{1/4}.$$

De wortels van deze vergelijking kunnen numeriek bepaald worden (zie o.a. Abramowitz, M. and Stegun, I.A., Handbook of math. functions. Dover publ., New York).
We vermelden de eerste 5 wortels in grootte:

(3.17)
$$\begin{array}{ll} \nu_1 = 1.8751041 & \nu_4 = 10.9955407 \\ \nu_2 = 4.6940911 & \nu_5 = 14.1371684 \\ \nu_3 = 7.8547574 & \nu_6 = \tfrac{1}{2}(2n-1)\pi,\ n \geq 6. \end{array}$$

Als coördinaatfuncties kiezen we het stelsel $\{x^k \mid k \geq 2\}$ dat volledig is in de ruimte $H \cap \{u \mid u(0) = u'(0) = 0\}$ met inprodukt (3.15). Vervolgens berekenen we de inprodukten $(A\phi_k,\phi_l)$ en (ϕ_k,ϕ_l) voor $k, l \geq 2$. De volgende betrekkingen worden afgeleid hiervoor:

(3.18)
$$\begin{cases} (A\phi_k,\phi_l) = \displaystyle\int_0^1 \frac{d^2x^k}{dx^2}\frac{d^2x^l}{dx^2}\,dx = \frac{k(k-1)l(l-1)}{k+l-3}, & k,l \geq 2 \\ (\phi_k,\phi_l) = \displaystyle\int_0^1 x^{k+l}\,dx = \frac{1}{k+l+1}, & k,l \geq 2. \end{cases}$$

Tenslotte lossen we het stelsel $(B - \lambda C)\vec{\alpha} = 0$ met $B_{kl} = (A\phi_k,\phi_l)$ en $C_{kl} = (\phi_k,\phi_l)$ voor $k, l \geq 2$ op. De resultaten staan weer in tabelvorm.

n	$\lambda_1^{(n)}$	$\lambda_2^{(n)}$	$\lambda_3^{(n)}$	$\lambda_4^{(n)}$	$\lambda_5^{(n)}$
1	20				
2	12.48	1211.5198			
3	12.3698	494.322	13958		
4	12.36240	490.969	4012.8	79296	
5	12.36298	485.515	3999.3	16517	316641
6	12.37044	485.396	3812.9	16489	49954
$\lambda_k = \nu_k^4$	12.362364	485.5188	3806.546	14617.3	39943.8

Als we deze resultaten nader bekijken, zien we voor n = 5 al een verslechtering optreden voor de waarde van λ_1; voor de andere eigenwaarden treedt dit moment iets later op. De volgende tabel geeft de relatieve fout waarmee $\lambda_k^{(n)}$ berekend is voor toenemende n.

n	$\lambda_1^{(n)}$	$\lambda_2^{(n)}$	$\lambda_3^{(n)}$	$\lambda_4^{(n)}$	$\lambda_5^{(n)}$
1	$.62_{10}{-0}$				
2	$.95_{10}{-2}$	$.15_{10}{+1}$			
3	$.60_{10}{-3}$	$.18_{10}{-1}$	$.27_{10}{+2}$		
4	$.30_{10}{-5}$	$.11_{10}{-1}$	$.54_{10}{-1}$	$.44_{10}{+1}$	
5	$.50_{10}{-4}$	$.34_{10}{-4}$	$.51_{10}{-1}$	$.13_{10}{-0}$	$.69_{10}{+1}$
6	$.65_{10}{-3}$	$.25_{10}{-3}$	$.17_{10}{-2}$	$.13_{10}{-0}$	$.25_{10}{-0}$

3.4. Ondergrenzen van eigenwaarden

In de vorige paragraaf is het construeren van bovengrenzen van eigenwaarden van een K-p.d. operator aan de orde geweest. Het probleem van het bepalen van ondergrenzen is in het algemeen erg moeilijk op te lossen; omdat het in de praktijk vaak om differentiaaloperatoren gaat die een randwaardeprobleem beschrijven, zullen we in het kort een methode aangeven om deze ondergrenzen te bepalen. Het randwaardeprobleem wordt m.b.v. de bijbehorende Greense functie overgevoerd in een integraalvergelijking. En voor de integraaloperator kunnen we nu de eigenwaarden gaan bepalen.
We zullen eerst enkele begrippen invoeren en een tweetal stellingen.

Zij $u \in L_2(a,b)$, dan beschouwen we het probleem:

$$(3.19)\quad \begin{cases} u(x) - \lambda \int_a^b K(x,s)u(s)ds = 0 \text{ met} \\ 1^\circ.\ K(x,s) = \overline{K(s,x)} \text{ (symmetrische kern) en} \\ 2^\circ.\ \int_a^b |K(x,s)|^2 ds \text{ is eindig.} \end{cases}$$

Onder de m-de geïtereerde kern $K_m(x,s)$ m.b.t. de kern $K(x,s)$ verstaan we

$$(3.20)\qquad \begin{cases} K_m(x,s) = \int_a^b K_r(x,t)K_{m-r}(t,s)dt & \text{voor iedere } r < m, \\ K_1(x,s) = K(x,s). \end{cases}$$

Uit de symmetrie van $K(x,s)$ volgt onmiddellijk de symmetrie van $K_m(x,s)$ voor $m \geq 1$. Als $K(x,s)$ een symmetrische kern is, niet identiek nul, dan geldt:

1°. K heeft ten minste één eigenwaarde;
2°. iedere eigenwaarde van K is reëel;
3°. eigenfuncties van verschillende eigenwaarden zijn orthogonaal.

Nu volgen de twee stellingen die we nodig hebben:

<u>stelling 3.4</u>

Zij $K(x,s)$ een symmetrische kern die voldoet aan $\int_a^b |K(x,s)|^2 ds < C$; laat verder $\{\phi_k\}$ de rij van de eigenfuncties van K zijn en $\{\lambda_k\}$ de rij van de bijbehorende eigenwaarden. Dan convergeert $\sum_{k=1}^{\infty} \frac{|\phi_k(x)|^2}{\lambda_k^2}$, en is deze som kleiner of gelijk aan de constante C.

bewijs: zie [5].

<u>stelling 3.5 (Hilbert-Schmidt)</u>

Laten $\{\lambda_k\}$ eigenwaarden zijn van een symmetrische kern $K(x,s)$ en $\{\phi_k\}$ de bijbehorende eigenfuncties. Zij $h = h(x)$ een functie waarvoor geldt dat $h^2(x)$ absoluut integreerbaar is over (a,b). Als de integraal $\int_a^b |K(x,s)|^2 ds$ begrensd is, dan kan de functie $f(x) = \int_a^b K(x,s)h(s)ds$ ontwikkeld worden in een absoluut en uniform convergente Fourrierreeks m.b.t. het orthonormale stelsel $\{\phi_k\}$:

$$f(x) = \sum_{k=1}^{\infty} f_k\phi_k, \; f_k = (f,\phi_k), \; k \geq 1.$$

De Fourier-coëfficiënten f_k van $f(x)$ voldoen tevens aan de volgende betrekking:

$$f_k = h_k/\lambda_k, \; h_k = (h,\phi_k), \; k \geq 1, \text{ zodat geldt: } f(x) = \sum_{k=1}^{\infty} \frac{h_k}{\lambda_k} \phi_k.$$

Voor het bewijs van deze stelling zie o.a. [5].

De stelling van Hilbert-Schmidt passen we toe voor $h(x) = K(x,t)$. Dan is aan de voorwaarden van deze stelling voldaan. Er volgt:

$$K_2(x,t) = \int_a^b K(x,s)K(s,t)ds = \sum_{k=1}^{\infty} \frac{\omega_k(t)}{\lambda_k} \phi_k(t), \text{ met } \omega_k(t) = \int_a^b K(x,t)\overline{\phi_k(x)}dk$$

$$(k \geq 1).$$

Omdat K symmetrisch is en $\phi_k(t)$ aan (3.19) voldoet met $\lambda = \lambda_k$, leiden we hieruit af, dat $\omega_k(t) = \frac{1}{\lambda_k} \overline{\phi_k(t)}$, voor $k \geq 1$. Als we dit substitueren, dan krijgen we de relatie:

$$K_2(x,t) = \sum_{k=1}^{\infty} \frac{\phi_k(x)\overline{\phi_k(t)}}{\lambda_k^2} .$$

Met volledige inductie naar m leiden we af:

$$(3.21) \qquad K_m(x,t) = \sum_{k=1}^{\infty} \frac{\phi_k(x)\,\overline{\phi_k(t)}}{\lambda_k^m} .$$

Deze reeks convergeert uniform m.b.t. x en t op [a,b] voor $m \geq 3$:

$$\left| \sum_{k=n+1}^{n+p} \frac{\phi_k(x)\phi_k(t)}{\lambda_k^m} \right| \leq \frac{1}{|\lambda_{n+1}^{m-2}|} \sum_{k=n+1}^{\infty} \left|\frac{\phi_k(x)}{\lambda_k}\right| . \left|\frac{\phi_k(t)}{\lambda_k}\right| \leq \frac{1}{|\lambda_{n+1}^{m-2}|} \left(\sum_{k=1}^{\infty} \frac{|\phi_k(x)|^2}{\lambda_k^2} \right)^{\frac{1}{2}} *$$

$$\left(\sum_{k=1}^{\infty} \frac{|\phi_k(t)|^2}{\lambda_k^2} \right)^{\frac{1}{2}} \leq \frac{C}{|\lambda_{n+1}^{m-2}|} .$$

De reeks (3.21) heet de bilineaire reeks voor $K_m(x,t)$. Onder het m-de spoor van een kern $K(x,s)$ verstaan we de integraal

(3.22) $$a_m = \int_a^b K_m(x,x)dx.$$

Als K symmetrisch is, en de rij $\{\phi_k\}$ van eigenfuncties van K is orthonormaal, dan leiden we uit de definitie van spoor af dat:

(3.23) $$a_m = \int_a^b \left(\sum_{k=1}^{\infty} \frac{|\phi_k(x)|^2}{\lambda_k^m} \right) dx = \sum_{k=1}^{\infty} \frac{1}{\lambda_k^m} \int_a^b |\phi_k(x)|^2 dx = \sum_{k=1}^{\infty} \frac{1}{\lambda_k^m}, \text{ en:}$$

(3.24) $$a_{2m} = \sum_{k=1}^{\infty} \frac{1}{\lambda_k^{2m}} > 0$$ (omdat λ_k reëel is, $k \geq 1$); verder geldt nog:

(3.25) $$a_{2m} = \int_a^b \int_a^b |K_m(x,t)|^2 dtdx.$$

Uit (3.25) zien we, dat voor de berekening van a_{2m} alleen K_m gekend moet worden; verder blijkt uit (3.24) dat we de eigenwaarden van K met het spoor a_m kunnen benaderen, omdat geldt: $\lambda_k^{2m} \leq \frac{1}{a_{2m}}$ voor $k \geq 1$.
Er geldt:

(3.26) Als λ_1 de kleinste eigenwaarde is van de symmetrische kern $K(x,t)$, dan geldt:

$$\lim_{n\to\infty} \sqrt[m]{\frac{1}{a_{2m}}} = \lim_{n\to\infty} \frac{a_{2m}}{a_{2m+2}} = \lambda_1^2.$$

<u>bewijs</u>:
stel λ_1 heeft multipliciteit p, $-\lambda_1$ multipliciteit q. Dan is de multipliciteit van λ_1^2 gelijk aan $r = p+q$. Zij verder:

$$a_{2m} = \frac{r}{2m}(1+\varepsilon_m) \text{ met } \varepsilon_m = \frac{1}{r} \sum_{k=r+1}^{\infty} \left(\frac{1}{\lambda_k}\right)^{2m} = 0.$$

We bewijzen dat $\varepsilon_m \to 0$ voor $m \to \infty$.
Stel in ε_m komt de eerste term n-keer voor. Voor $k > r$ geldt $|\lambda_k| > |\lambda_1|$, dus:

$$\varepsilon_m = \frac{1}{r}\left(\frac{\lambda_1}{\lambda_{r+1}}\right)^{2m} \left[n + \sum_{k=r+n+1}^{\infty} \left(\frac{\lambda_{r+1}}{\lambda_k}\right)^{2m} \right] \leq \frac{1}{r}\left(\frac{\lambda_1}{\lambda_{r+1}}\right) \left[n + \sum_{k=1}^{\infty} \left(\frac{\lambda_{r+1}}{\lambda_k}\right)^2 \right] \to 0$$ voor $m \to \infty$, omdat de laatste som begrensd is.

Uit dit bewijs leiden we nog de volgende praktische betrekkingen af:

(3.27a) $$|\lambda_1| = \sqrt[2m]{\frac{r(1+\varepsilon_m)}{a_{2m}}} > \sqrt[2m]{\frac{r}{a_{2m}}}$$, benedengrens voor λ_1;

(3.27b) $$\frac{a_{2m}}{a_{2m+2}} = \lambda_1^2 \frac{1+\varepsilon_m}{1+\varepsilon_{m+1}} > \lambda_1^2,$$ bovengrens voor λ_1(want $\varepsilon_m > \varepsilon_{m+1}$).

Niet alleen voor de kleinste eigenwaarde, maar ook voor de andere kunnen dergelijke relaties afgeleid worden. Voor de op een na kleinste eigenwaarde λ_2, aannemende dat λ_1 en λ_2 enkelvoudig zijn, leiden we af:

(3.28) $$b_{2m} = a_{2m}^2 - a_{4m} = \left(\sum_{k=1}^{\infty} \frac{1}{\lambda_k^{2m}}\right)^2 - \sum_{k=1}^{\infty} \frac{1}{\lambda_k^{4m}} =$$

$$= 2 \sum_{l=2}^{\infty} \sum_{k=1}^{\infty} \frac{1}{(\lambda_k\lambda_l)^{2m}} \geq \frac{2}{(\lambda_1\lambda_2)^{2m}} \Longrightarrow |\lambda_2| \geq \frac{1}{|\lambda_1|} \sqrt[2m]{\frac{2}{b_{2m}}} .$$

Als λ_1 bekend is, is dit een ondergrens voor λ_2.

We geven nu een voorbeeld. Daarvoor is gekozen het Sturm-Liouville-probleem voor de snaar. Het is bekend dat het probleem $Lu = \lambda u$ met $u \in G$, G een gebied, en $u = 0$ op ∂G, voor het geval dat $\lambda = 0$ geen eigenwaarde is van L, equivalent is met $u = \lambda K u$, waarin K de Greense functie is van het oorspronkelijke probleem.

<u>voorbeeld</u>

(3.29) $$\begin{cases} -\dfrac{d^2u}{dx^2} = \lambda u \text{ in } (0,\pi), \\ \\ u(0) = u(\pi) = 0, \text{ en } u \in L_2(0,1). \end{cases}$$

De eigenfuncties van $-\dfrac{d^2}{dx^2}$ zijn in dit geval $\phi_k = \sin(kx)$, $(k \geq 1)$,en de eigenwaarden $\lambda_k = k^2$, $(k \geq 1)$.

De Greense functie die bij dit probleem hoort, wordt gegeven door:

(3.30) $$\begin{cases} G(x,s) = x(1 - \frac{s}{\pi}), \ x < s \text{ en} \\ \\ = s(1 - \frac{x}{\pi}), \ x > s. \end{cases}$$

(3.29) is nu equivalent met

$$(3.31) \qquad u(x) = \lambda \int_0^{\pi} G(x,s)u(s)ds.$$

We moeten de tweede geïtereerde kern bepalen.
$G_1(x,s) = G(x,s)$, en $G_2(x,s) = \int_0^{\pi} G_1(x,t)G(t,s)dt$.
Voor $x > s$ geldt:

$$(3.32a) \qquad G_2(x,s) = \int_0^{\pi} G(x,t)G(t,s)dt = \int_0^{s} t^2(1 - \frac{x}{\pi})(1 - \frac{s}{\pi})dt +$$

$$\int_s^{x} s(1 - \frac{t}{\pi})t(1 - \frac{x}{\pi})dt + \int_x^{\pi} sx(1 - \frac{t}{\pi})^2 dt =$$

$$\frac{1}{6}\left[- s^3(1 - \frac{x}{\pi}) + s(\frac{x^3}{\pi} - 3x^2 + 2\pi x)\right].$$

Wegens de symmetrie van $G_2(x,s)$ geldt voor $x < s$:

$$(3.32b) \qquad G_2(x,s) = \frac{1}{6}\left[- x^3(1 - \frac{s}{\pi}) + x(\frac{s^3}{\pi} - 3s^2 + 2\pi s)\right].$$

Hieruit berekenen we a_2 en a_4.

$$(3.33) \qquad \begin{aligned} a_2 &= \int_0^{\pi}\int_0^{\pi} |G_1(x,s)|^2 dsdx = 2\int_0^{\pi} dx \int_0^{x} s^2(1 - \frac{x}{\pi})^2 ds = \frac{\pi^4}{90}. \\ a_4 &= \int_0^{\pi}\int_0^{\pi} |G_2(x,s)|^2 dsdx = \frac{\pi^8}{9450}. \end{aligned}$$

We berekenen uit deze gegevens grenzen voor λ_1 en λ_2 m.b.v. de formules (3.27a), (3.27b), en (3.28).
De resultaten staan in de volgende tabel vermeld:

m	$\lambda_1^{(m)}$	rel. fout	$\lambda_2^{(m)}$	rel. fout
1	.9612	$.93_{10}-1$		
2	.9989	$.11_{10}-2$	3.46	$.14_{10}-0$
$\lambda_k=k^2$	1		4	

Dit zijn ondergrenzen voor λ_1 en λ_2. Verder geldt nog de bovengrens:
$\lambda_1 \leq 1.038$, met een rel.fout van $.38_{10}-1$.

Literatuur:

[1] Dzhishkariani, A.V., The least squares and Bubnov-Galerkin methods. U.S.S.R. Comp. Math. and Math. Phys., vol. 8, Pergamon press, Oxford (1970).

[2] Gould, S.H., Variational methods for eigenvalueproblems. University of Toronto press, Math. exp. no. 10, Toronto (1957).

[3] Kaczmarz, S. en Steinhaus, H., Theorie der Orthogonalreihen. Chelsea publ. Co., New York (1951).

[4] Lauwerier, H.A., Calculus of variations in mathematical physics. MC tract 14, Mathematisch Centrum, Amsterdam (1966).

[5] Mikhlin, S.G., Integral equations. Int.ser. of mon. on Pure and Applied Math., vol. 4, Pergamon press, London (1957).

[6] Mikhlin, S.G., The problem of the minimum of a quadratic functional. Holden-Day ser. in math. phys., London (1965).

[7] Mikhlin, S.G., Variationsmethoden der mathematischen Physik. Akademie-Verlag, Berlijn (1962).

[8] Mikhlin, S.G., The numerical performance of variational methods. Wolters-Noordhoff Publ., Groningen (1971).

[9] Petryshyn, W.V., Direct and iterative methods for the solution of linear operator equations in Hilbert space. Trans.Am.Math.Soc., vol. 105, (1962).

[10] Riesz, F. en Sz.-Nagy, B., Leçons d'analyse fonctionnelle. Académie des Sciences de Hongrie, Budapest (1952).