Nella presente versione 2017 sono state apportate alcune modifiche alla precedente versione 2014 riveduta e ampliata.

Desidero ringraziare calorosamente due allievi di qualche anno fa: il dott.Vincenzo Parente, che con notevole pazienza ha effettuato la scansione di numerose pagine manoscritte nonché il ritocco di quelle deteriorate, e l'ing. Mauro Russo, che ha eseguito tecnicamente le correzioni e nel contempo, soffermandosi sul testo con interesse e passione, ha rilevato numerose e pertinenti osservazioni che hanno contribuito al miglioramento della precedente versione.

Ritengo doveroso evidenziare che all'ing. Russo è dovuta la dimostrazione dell'equivalenza tra norme di pag.104, che era stata esposta solo nel caso m=1, nonché la proposizione (8.1) di pag.182.5 .

Renato Fiorenza

Università degli Studi di Napoli Federico II
Scuola di Dottorato in Ingegneria Industriale

APPUNTI
delle lezioni di
ANALISI FUNZIONALE

versione 2017 riveduta e ampliata

Renato Fiorenza

Questo volume è stato realizzato dal COINOR
Centro di Ateneo per la Comunicazione e
l'innovazione Organizzativa
Università degli Studi di Napoli Federico II
Corso Umberto I, n. 40
80138 Napoli
www.coinor.unina.it
TUTTI I DIRITTI RISERVATI

Responsabile del progetto editoriale Giuseppe Zollo
Coordinamento operativo Giovanni Miano
Progetto grafico e impaginazione Maria Grazia Mergenni
Stampa Print Agency by I Farella Snc

© 2005 - Università degli Studi di Napoli Federico II

ISBN 88-7893-003-2

Con il contributo di Ministero dell'Istruzione, dell'Università e della Ricerca, Centro Ricerche FIAT, Corriere del Mezzogiorno, Comune di Napoli, EdiSES, Fondazione Istituto Banco di Napoli, Labozeta, Provincia di Napoli, Regione Campania, Sanpaolo Banco di Napoli.

Renato Fiorenza

APPUNTI DELLE LEZIONI DI
ANALISI FUNZIONALE

versione 2017 riveduta e ampliata

Scuola di Dottorato in Ingegneria Industriale

GLI STRUMENTI DI COINOR

PRESENTAZIONE

di Guido Trombetti

Le discipline matematiche sono uno strumento prezioso per affrontare molti problemi nell'ambito delle scienze applicate. In particolare, sulle nozioni fondamentali di analisi funzionale si reggono gran parte dei modelli matematici che consentono di rappresentare efficacemente i fenomeni della realtà.

D'altra parte, lo sviluppo delle teorie matematiche tende naturalmente alla sintesi e all'unificazione intorno a semplici idee di base. Un'esposizione di tali teorie che sia a sua volta rigorosa e rispettosa della sintesi cui si è pervenuti, darà sicuramente l'impressione di essere arrivati ad un punto di perfezione, poiché essa sarà completa, elegante e breve. Ma può un docente di matematica seguire questa naturale propensione quando vuole comunicare le sue conoscenze ad una platea costituita da futuri ingegneri, economisti, biologi, etc.?

Una volta Henri Lebesgue disse: "Les mathématiques ont été créés par les hommes pur leurs besoins et elles sont en fait un auxiliaire précieux; le professeur de mathématiques doit rester un professeur d'action". Senza dubbio il docente di matematica, se vuole far crescere la capacità di penetrazione del proprio insegnamento, non può limitarsi a comunicare il proprio sapere così come lo conosce, pretendendo che gli altri si adattino. Deve invece utilizzare una esposizione che, senza perdere rigore, tenga presente il livello delle conoscenze degli allievi e la terminologia adottata nelle discipline collegate.

Il corso di Analisi Funzionale tenuto da Renato Fiorenza per circa trenta anni presso la Facoltà di Ingegneria dell'Università di Napoli Federico II è sicuramente un esempio di didattica della matematica improntata alle idee appena esposte. Non a caso, per ammissione generale, esso è considerato un importante strumento per acquisire i concetti matematici alla base dei metodi di interesse per gli ingegneri. Gli appunti relativi a tale corso sono stati preparati sin dal primo anno ed hanno subito nel tempo cambiamenti volti ad arricchirne i contenuti, soprattutto con l'obiettivo di ampliare lo spazio dedicato ad esempi.

La lodevole ed apprezzata iniziativa del Consiglio Scientifico della Scuola di Dottorato in Ingegneria Industriale ha fornito all'Ateneo Federico II l'occasione per pubblicare in una forma completa questi Appunti. L'opera non si presenta come un tradizionale manuale di Analisi Funzionale.

E' abbastanza facile trovare libri che per larga parte sono dedicati ad enunciati e dimostrazioni di teoremi, ma con poco spazio dedicato a mettere in luce le possibili applicazioni. Negli Appunti di Renato Fiorenza, invece, fin dai primi capitoli, vengono introdotte tutte le strutture che saranno utilizzate (spazi topologici, metrici, normati, etc.) e, facendo riferimento ad alcuni significativi esempi di spazi funzionali, in particolare quelli di Lebesgue e di Sobolev, si mettono in evidenza le peculiarità di ciascuna di esse. Anche lo spettro delle informazioni reperibili è un po' più ampio di quello tradizionalmente coperto dai testi e va sottolineata l'attenzione rivolta ai problemi differenziali. Per un elenco dettagliato degli argomenti trattati si fa riferimento all'indice che con la sua organizzazione a due livelli, dapprima un livello schematico e quindi uno più particolareggiato, faciliterà l'orientamento del lettore.

Sono sicuro che questo volume continuerà ad essere punto di riferimento sia come strumento didattico per affrontare corsi con contenuti di analisi funzionale, sia come strumento di consultazione per quanti, anche non specialisti della materia, vogliano trovare rapidamente informazioni sui risultati di interesse.

A Renato Fiorenza un mio personale ringraziamento per la passione e la raffinata professionalità con la quale ha svolto, e svolge tuttora, un'opera di diffusione della cultura matematica di rara efficacia.

PREFAZIONE

di Renato Fiorenza

L' obiettivo del corso di Analisi Funzionale che ho tenuto nella facoltà di Ingegneria dal 1975 per circa trenta anni, era quello di presentare, senza troppi dettagli, una panoramica dei concetti e dei principali risultati atta a sviluppare, nell' allievo da avviare all' attività scientifica in uno dei settori dell'Ingegneria, la capacità di ricercare ed approfondire nei testi di Analisi Funzionale l' argomento che interessa ai suoi scopi, o addirittura il risultato che può essergli utile.

Perciò, piuttosto che dedicare le lezioni ad una particolare struttura topologica (ad es. spazio di Hilbert), ho preferito introdurre appena possibile le varie strutture, ed illustrare concetti e risultati evidenziando di volta in volta il loro ruolo. Per rendere l' idea, tale impostazione consente, dopo aver introdotto la convergenza di una successione, o la continuità di una funzione, di sottolineare le differenti conseguenze secondo che l'ambiente sia uno spazio vettoriale topologico, o uno spazio metrico, o uno spazio normato, o uno spazio di Hilbert: ad esempio consente di far notare che, mentre la continuità di una funzione definita in uno spazio metrico equivale alla continuità di essa "per successioni", l' equivalenza non sussiste con riferimento ad una funzione definita in uno spazio topologico qualunque.

Ho introdotto fin dall' inizio vari spazi funzionali, presentandoli come esempi di spazi vettoriali; successivamente, al momento opportuno, a ciascuno è stata associata la topologia usuale, sottolineando che non sempre essa è generata da una norma. Per ciascuno spazio è stato specificato il significato della convergenza di una successione: esso ovviamente si riflette sul significato, importante per un allievo ingegnere, del termine "approssimazione" riferito alla soluzione di un problema. Quanto sopra consente poi di presentare esempi di operatori lineari tra spazi funzionali diversi, e di illustrarne la continuità.

Piuttosto che soffermarmi sulle dimostrazioni, ho preferito dare risalto ai concetti, e soprattutto convincere l' allievo dell' utilità, anche per le applicazioni, di certe nozioni astratte e di certi teoremi molto generali. E' a questo scopo, ad esempio, che ho mostrato come la teoria astratta delle disequazioni variazionali si applica allo studio della configurazione di equilibrio di un filo flessibile, fissato agli estremi e teso su un ostacolo, o di una trave incastrata e soggetta a carichi concentrati; l' allievo viene anche a constatare che la costruzione di un modello matematico, adatto alla questione, richiede la conoscenza non solo di certe nozioni di Analisi Funzionale, ma anche e soprattutto di certi spazi funzionali.

Avendo voluto non entrare nei dettagli, è stato possibile parlare ad esempio di soluzioni deboli delle equazioni differenziali, del problema di Dirichlet in forma debole, di serie di Fourier in uno spazio di Hilbert, di distribuzioni (in particolare quella di Dirac), di equazioni integrali, di teoria spettrale, della rappresentazione spettrale di un operatore lineare (compatto e autoaggiunto), di problemi di punto fisso, di calcolo delle variazioni: come già detto, tutto in forma "panoramica".

* * *

Ho voluto accennare alla linea seguita nell' impostazione del corso per dare così un' idea di questi Appunti, che costituiscono un' esposizione scritta delle lezioni; sono stati stampati in seguito alla delibera del Consiglio Scientifico della Scuola di Dottorato in Ingegneria Industriale, che ha istituito il corso di Analisi Funzionale. La stampa è stata realizzata sulla base di una scansione delle fotocopie degli appunti, che erano state distribuite agli studenti negli ultimi anni in cui ho svolto il corso. La stesura è stata effettuata nel 1975, ma quella attuale comprende diverse aggiunte e varie modifiche apportate negli anni successivi.

In un primo momento le pagine erano state dattiloscritte e poi ciclostilate; in seguito sono state fotocopiate e, gradualmente negli anni, sono state modificate inserendo delle correzioni manoscritte (realizzate con foglietti incollati sulle fotocopie). Sono state anche aggiunte molte pagine nuove.

Per non modificare la numerazione delle pagine, che interviene in tutti i riferimenti, è stato necessario numerare in modo autonomo quelle nuove, e talvolta creare una numerazione “doppia”, o eliminare qualche pagina (che perciò risulta mancante).

Il materiale è organizzato in modo da consentire uno studio su piani di approfondimento diversi ; ad esempio le dimostrazioni sono comprese fra due linee tratteggiate e quindi possono essere omesse in una prima lettura.

Il Cap. I è preceduto da alcune premesse, che constano di diciannove pagine indicate con i numeri racchiusi in un quadratino; dopo l’ ultimo capitolo è stata inserita un’ appendice, che comprende non solo qualche complemento ad argomenti esposti in precedenza, ma anche alcuni appunti che avevo scritto per altre occasioni, indipendenti dal corso di Analisi Funzionale.

Per ottenere una stesura ben organizzata occorreva effettuare una nuova esposizione, utilizzando quella attuale come bozza; mi è sembrato un lavoro eccessivo in rapporto al contenuto, che non è, e non voleva essere, quello di un libro di testo.

Siccome l’ indice è stato predisposto con molti dettagli e perciò consta di undici pagine (numerate $\langle 1 \rangle$, $\langle 2 \rangle$, ...), ho ritenuto opportuno farlo precedere da un “ indice schematico “, che per ciascun capitolo indica la pagina dell’ indice in cui esso si trova. E’ stato anche realizzato l’ indice analitico.

Desidero esprimere il mio vivo ringraziamento al rettore Trombetti per aver consentito, cosa che mi ha molto lusingato, la stampa degli Appunti a cura dell’ Università degli Studi Federico II, accogliendo la richiesta del Consiglio Scientifico della Scuola di Dottorato in Ingegneria Industriale.

Ringrazio anche l’ allievo Giovanni Di Fratta per gli schemi a colori delle pagine 48bis e 48ter, che ha realizzato durante il corso dimostrando impegno e capacità.

Ringrazio altresì cordialmente l’ allievo Pierpaolo Tamasi, che mi ha fornito la scansione delle fotocopie con la relativa impaginazione: ha effettuato il lavoro anche eseguendo un accurato e paziente ritocco, con encomiabile zelo e di sua iniziativa dopo aver superato brillantemente l’ esame.

INDICE SCHEMATICO

Cliccando sul titolo o sul numero di pagina si passa all'indice dettagliato e da questo, analogamente, si passa all'argomento cercato. Per tornare a questa pagina digitare I *; dall'indice dettagliato si può anche cliccare in alto sul numero della pagina.*

Premesse ……..…..…..⟨1⟩

Capitolo I - Spazi topologici .…….……⟨1⟩

Capitolo II - Cenni sulla teoria della misura e dell'integrazione secondo Lebesgue. Cenni sull'integrale di Stieltjes .....⟨2⟩

Capitolo III - Alcuni richiami di Algebra. Operatori e funzionali lineari ….……..⟨2⟩

Capitolo IV - Spazi metrici. Spazi vettoriali topologici. Spazi vettoriali normati. Spazi pre-hilbertiani ……....⟨4⟩

Capitolo V - Successioni e serie. Spazi metrici completi. Spazi di Banach. Spazi di Hilbert …….....⟨5⟩

Capitolo VI - Operatori lineari e continui tra spazi vettoriali normati. Funzionali lineari e continui …….....⟨5⟩

Capitolo VII - Lo spazio degli operatori lineari e continui. Serie di operatori. Operatore aggiunto. Operatori autoaggiunti …⟨8⟩

Capitolo VIII - Cenni sulla teoria spettrale …….⟨9⟩

Capitolo IX - Disequazioni variazionali ed applicazioni …...…….⟨9⟩

Capitolo X - Cenni di Analisi non lineare …….⟨10⟩

Appendice ……⟨10⟩ , ⟨11⟩

INDICE ANALITICO……..⟨11⟩

INDICE

Cliccando sul titolo o sul numero di pagina si passa all'argomento cercato.
Cliccando in alto sul numero della pagina, si torna indietro o all' indice schematico.
Per accedere all' indice analitico digitare J, alle pagine
con quadratino digitare prima [ ] *e poi il numero della pagina.*

Premesse.

Relazioni di equivalenza. Insieme quoziente pag. 1
Insiemi equipotenti 8
Definizioni 12
La potenza di un insieme 14
Insiemi numerabili. Potenza del continuo 16

Capitolo I. Spazi topologici.

Introduzione pag. 1
1. Generalità 4
2. Sottospazi. Spazi di Hausdorff 9
3. Sottoinsiemi connessi di uno spazio topologico; sottoinsiemi compatti e sottoinsiemi precompatti 12
4. Assiomi di Hausdorff. Spazio topologico prodotto 16
5. Basi di topologia 18
6. Spazi separabili 19
7. Limite di una successione. Limite e continuità di una funzione in un punto 19
8. Confronto di topologie 24
Esempi ed osservazioni
1. Sulla topologia $\mathcal{P}(S)$ 27
2. Sulla topologia $\{\phi, S\}$ 27
3,4,5. Esempi 28, 29
6. Topologia della convergenza puntuale 30

Capitolo II. Cenni sulla teoria della misura e dell'integrazione secondo Lebesgue. Cenni sull'integrale di Stieltjes.

1. La misura secondo Lebesgue pag. 31
2. Funzioni misurabili 34
3. L'integrale secondo Lebesgue 37
4. Funzioni a valori in $\overline{R}$ definite quasi ovunque: sommabilità e integrale secondo Lebesgue 41
5. Proprietà dell'integrale secondo Lebesgue 43
6. La nozione di misura 44
7. Cenni sull'integrale di Stieltjes-Lebesgue 45
8. Funzioni a variazione limitata. Funzioni assolutamente continue 46
9. Integrale di Stieltjes di una funzione continua di una variabile 48

Capitolo III. Alcuni richiami di Algebra. Operatori e funzionali lineari.

1. Spazi vettoriali 49

Esempi di spazi vettoriali

1. Lo spazio $L^p(\Omega)$ 54
2. Lo spazio $L^\infty(\Omega)$ 54
3. Lo spazio $L^p(\Omega,\mu)$ 55
4. Lo spazio l_p 55
5. Lo spazio l_∞ 55
6, 7, 8 Esempi 55, 56
9. Lo spazio $C^0(T)$ 56
10. Lo spazio $C^m(T)$ 56

11. Lo spazio $C^\infty(T)$ 56
12. Lo spazio $C^{0,\lambda}(T)$ 56
13. Lo spazio $C^{m,\lambda}(T)$ 57
14. Lo spazio $C_0^\infty(T)$ 57
15. Lo spazio $C_0^m(\Omega)$ 57
16. Lo spazio $C_0^\infty(\Omega)$ 58
17. Lo spazio $L^p_{loc}(\Omega)$ 58
18. Derivate deboli. Lo spazio $W^{m,p}(\Omega)$ 59

2. Operatori lineari. Funzionali. Forme bilineari. Forme quadratiche 62

Esempi

1. Funzionali lineari definiti in R 64
2. Funzionali lineari definiti in R^k 64
3. Operatori lineari di R^k in R^m 64
4. Operatori lineari tra spazi a dimensione finita 65
5. Funzionali lineari definiti in $L^2(\Omega)$ 66
6. Esempio di operatore lineare di $L^p(\Omega)$ in $L^p(\Omega)$ 66
7. Esempio di funzionale lineare definito in $C_0^k(\Omega)$ 66
8. Esempio di operatore lineare di $C^2(T)$ in $C^0(T)$ 66
9. Esempio di funzionale lineare definito in $W^{m,p}(\Omega)$ 66

Forme bilineari. Forme quadratiche 67

Esempi

1. Forme bilineari in $C^k \times C^m$ 68
2. Esempi di forme bilineari in $W^1(\Omega)$ 68

3. Lo spazio degli operatori lineari di Σ in Σ'. Duale algebrico di uno spazio vettoriale 69

Esempi

1. Duale di R^K e di uno spazio a dimensione finita 70

2. Esempi sul duale algebrico di $L^p(\Omega)$ 70

4. Prodotto di operatori. Operatori lineari invertibili; inverso destro e inverso sinistro 72

Capitolo IV. Spazi metrici. Spazi vettoriali topologici. Spazi vettoriali normati. Spazi pre-hilbertiani.

1. Spazi metrici 77

2. Spazi vettoriali topologici 82

3. Spazi vettoriali normati 88

4. Spazi pre-hilbertiani 93

5. Alcuni spazi funzionali. Spazi di successioni 98

1) l_∞ ed $L^\infty(\Omega)$ 99

2) l_p ed $L^p(\Omega)$ 99

3) $L^p(\Omega,\mu)$ 100

4) $C^0(T)$ 100

5) $C^m(T)$ 100

6) $C^{0,\lambda}(T)$ 101

7) $C^{m,\lambda}(T)$ 101

8) $C^\infty(T)$ 102

9) $C_0^0(\Omega)$ 102

10) $C_0^m(\Omega)$ 102

11) $C_0^\infty(\Omega)$ 102

12) $L^p_{loc}(\Omega)$ 102

13) $W^{m,p}(\Omega)$ 103

14) Lo spazio di Sobolev $H^{m,p}(\Omega)$ 103

15) Lo spazio di Sobolev $H_0^{m,p}(\Omega)$ 103

Teorema di Sobolev 105

Capitolo V . Successioni e serie . Spazi metrici completi . Spazi di Banach . Spazi di Hilbert.

1. Successioni convergenti in uno spazio topologico 106

2. Sottoinsiemi sequenzialmente compatti di uno spazio topologico 111

3. Spazi metrici completi . Spazi di Banach. Spazi di Hilbert 112

4. Proiezione di un punto su un convesso chiuso di uno spazio di Hilbert. Proiezione ortogonale su un sottospazio chiuso. Proiettori 115

5. Sistemi ortonormali di punti di uno spazio pre-hilbertiano 122

6. Serie in uno spazio di Banach. Serie di Fourier in uno spazio di Hilbert 125

Esempi 131

Dimostrazione di alcuni risultati sullo spazio l_p 132.2

Capitolo VI . Operatori lineari e continui tra spazi vettoriali normati . Funzionali lineari e continui .

1. Applicazioni continue. Continuità uniforme. Isometrie 133

2. Operatori lineari e continui tra spazi vettoriali normati. Isomorfismi 137

3. Esempi 144

1. Continuità dei funzionali lineari definiti in R^k 144

2. Continuità degli operatori lineari di R^k in R^m 145

3. Continuità degli operatori lineari tra spazi a dimensione finita 145

4. Funzionali lineari e continui in $L^p(\Omega)$ 146

5. Continuita' dell'operatore differenziale lineare del 2° ordine a coefficienti continui 146

6. Esempio di funzionale lineare e continuo in $W^{m,p}(\Omega)$ 147

7. Esempio di operatore lineare e continuo di $H_0^m(\Omega)$ in $L^2(\Omega)$ 147

8. Continuita' dell'operatore differenziale del secondo ordine in forma di divergenza 147

9. Ulteriore esempio di operatore lineare e continuo di $H_0^2(\Omega)$ in $L^2(\Omega)$ 148

10. Esempio di funzionale lineare e continuo in $H_0^1(\Omega)$ 148

11. Esempio di funzionale lineare e continuo in $H_0^2(\Omega)$ 148

12. Esempio di funzionale lineare e continuo in $H_0^1(\Omega) \cap H^2(\Omega)$ 148

13. Distribuzioni su un aperto Ω 149

14. La distribuzione di Dirac 154, 157

15. Derivate di una distribuzione. Derivate nel senso delle distribuzioni. 158.2

4. Completamento di uno spazio vettoriale normato 159

5. Operatori compatti 165

Esempi

1) Equazioni integrali di Fredholm 168 (*)

(*) Cfr. anche 201, 203 bis, 218, 226, 236, 251.

2) Immersione di $C^{0,\lambda}(A)$ in $C^0(A)$ 170

3) Osservazione sull'identità in uno spazio di dimensione infinita 170

4) Osservazione sugli operatori compatti di uno spazio di dimensione infinita in sè 170

6. Duale topologico di uno spazio vettoriale topologico. Duale forte di uno spazio vettoriale normato 171

7. Duale di uno spazio di Hilbert 177, 179

8. Esempi 179

1) Duale di R^K 179

2) Duale di uno spazio a dimensione finita 70, 180

3) Duale di $L^2(\Omega)$ 180

4) Duale di $L^p(\Omega)$ 180

5) Duale di $L^\infty(\Omega)$ 182

6) Duale di $H^{m,p}(\Omega)$ 182

7) Esempi di elementi del duale di $C^0(\Omega)$ 182.1

8) Duale di $C_0^\infty(\Omega)$ 182.1

9) Un'applicazione del teorema di Riesz: problema omogeneo di Dirichlet generalizzato (o in forma debole) 182.2

10) Un'applicazione del problema di Dirichlet in forma debole 182.3

Note sui modelli matematici 182.4

9. Il teorema di Hahn-Banach. Alcuni corollari 183

10. Biduale di uno spazio vettoriale normato. Spazi riflessivi. Spazi di Banach uniformemente convessi 186

11. Topologia della convergenza puntuale nel duale di uno spazio vettoriale normato (topologia $*$-debole) 191

12. Topologia debole per uno spazio vettoriale normato 195

Capitolo VII . Lo spazio degli operatori lineari e continui - Serie di operatori. Operatore aggiunto. Operatori autoaggiunti.

1 . Lo spazio degli operatori lineari e continui di uno spazio normato in un altro . Topologia della convergenza forte e della convergenza debole degli operatori 202

Esempi

1. Norma di un operatore di R^k in R^m 203

2. Norma di un operatore tra spazi a dimensione finita 203 bis

3 . Norma dell'operatore di derivazione 203 bis

4 . Norma dell'operatore di derivazione in $H_0^m(\Omega)$ 203 bis

5 . Norma di un proiettore 203 bis

6 . Norma di un operatore differenziale lineare di $C^2(A)$ in $C^0(A)$ 203 bis

7. Norma di un operatore integrale 203 bis

2. Applicazioni bilineari continue . Forme sesquilineari continue 209

Forme bilineari coercitive 213

Esempi 213

3. Serie di operatori 215

Esempio sulle equazioni integrali di Fredholm 218

4. Aggiunto hermitiano 219

Una caratterizzazione dei proiettori 224

Esempi sulla nozione di operatore aggiunto

1. Aggiunto di un operatore lineare di C^k in C^k 227

2. Aggiunto di un operatore lineare di $H_0^m(\Omega)$ in $L^2(\Omega)$ 227.1

3-4. Aggiunto di alcuni operatori differenziali lineari in $H_0^2(\Omega)$ 227.2-227.3

5. Aggiunto di un operatore integrale 227.4

5. Aggiunto di un operatore lineare e continuo tra spazi normati 227.5

6. Il teorema dell'alternativa 233

Applicazione alle equazioni integrali di Fredholm 236

Esercizio sulle equazioni integrali 237

Capitolo VIII. Cenni sulla teoria spettrale.

1. Insieme risolvente e spettro di un operatore lineare di uno spazio vettoriale topologico in sé. Autovalori, autovettori 239

2. Insieme risolvente e spettro di un operatore lineare e continuo di uno spazio di Banach in sé. Raggio spettrale 240

3. Caso di uno spazio di Hilbert 242

4. Rappresentazione spettrale di un operatore compatto e autoaggiunto in uno spazio di Hilbert. Operatori del tipo $f(A)$, con f funzione reale ed A operatore 245

Applicazione alle equazioni integrali di Fredholm 251

Capitolo IX. Disequazioni variazionali ed applicazioni.

1. Disequazioni variazionali 254

2. Formulazione variazionale di un particolare problema ai limiti 258

3. Un problema sul filo flessibile teso su un ostacolo rigido 260

4. Un problema sul filo flessibile teso su un ostacolo deformabile 267

5. Un problema sulla trave incastrata 274

6. Un problema sulla trave incastrata e forzata su un ostacolo 278

Capitolo X . Cenni di Analisi non lineare.

1. Punti fissi 284
2. Differenziale 287
Esempi
1. Differenziale di una funzione vettoriale di n variabili reali 295
2. L'integrale del Calcolo delle variazioni e l'equazione di Eulero 295 bis
3. Differenziali di ordine superiore 298
4. Cenni sull'integrazione di una funzione continua a valori in uno spazio di Banach 301

Appendice.

Degli appunti qui raccolti, alcuni costituiscono dei complementi a questioni esposte in precedenza, altri sono stati redatti in occasioni diverse, indipendentemente dal corso di Analisi Funzionale.

— • —

- Teorema di Sobolev 304
- Misure spettrali in uno spazio di Hilbert. Rappresentazione spettrale di un operatore autoaggiunto 306
- Perturbazioni delle disequazioni variazionali 312
- Un teorema astratto di regolarità per le soluzioni delle disequazioni variazionali 317
- Sottodifferenziale 326
- Un'osservazione sul teorema di Vitali - Lusin 329
- Confronto tra le nozioni: hölderianità - assoluta continuità - continuità uniforme 330
- Descrizione degli argomenti esposti successivamente, con numerazione autonoma delle pagine 335

⟨11⟩

Le pagine che seguono sono indicate con numeri preceduti da una lettera maiuscola e da un punto. Le stesse notazioni vanno digitate per accedervi da altre pagine.

DOPO LA PAGINA 335

1) Introduzione alle equazioni integrali di Fredholm A.1-A.16

2) Operatori integrali. Equazioni integrali di Fredholm B.1-B.16

3) Forme differenziali di grado K C.1-C.13

4) Integrale superficiale di una K-forma C.14-C.20

5) Una funzione continua in un intervallo, con derivata nulla quasi ovunque, che non è costante D.1-D.7

INDICE ANALITICO J

Premesse

1

RELAZIONI DI EQUIVALENZA

INSIEME QUOZIENTE

Pensiamo ad un giocatore di roulette che punta sul pari: per lui è indifferente se la pallina si ferma sul 2 o sul 4, sul 6 o sul 10; cioè per quel giocatore i numeri 2,4,6,... sono da considerarsi tutti uguali, e così pure i numeri 1,3,5,...

Pensiamo di dover calcolare il seno e il coseno di assegnati numeri reali. A questo scopo i numeri x, $x-2\pi$, $x+2\pi$, $x-4\pi$, $x+4\pi$ ecc. si comportano tutti allo stesso modo, giacché hanno tutti lo stesso seno e lo stesso coseno; di conseguenza, per lo scopo suddetto, i numeri del tipo $x+2k\pi$ con $x \in Z$ sono da considerarsi tutti uguali.

Pensiamo ad un periodo nel quale, essendo stata emessa una nuova banconota da dieci euro, sono in circolazione due tipi di banconote da dieci: il cittadino può adoperare indifferentemente la nuova o la vecchia banconota, e perciò esse sono da considerarsi uguali.

I precedenti esempi mostrano che talvolta in un insieme S si può avere l'esigenza, o l'opportunità, di considerare certe coppie di elementi distinti come se fossero coppie di elementi uguali; in termini più precisi, si può avere l'esigenza o l'opportunità di considerare "come uguaglianza" un'altra "relazione" che svolge, per la questione che si vuole affrontare, lo stesso ruolo della relazione di uguaglianza. Nasce così il concetto di "relazione di equivalenza".

Dato un insieme S, denotiamo con $\mathcal{R}$ una proprietà definita nel prodotto cartesiano $S \times S$, sicché una coppia ordinata (x,y) di elementi di S può godere o non della proprietà $\mathcal{R}$: in caso affermativo si scrive $x\mathcal{R}y$, mentre la scrittura $x\not\mathcal{R}y$ esprime che la coppia ordinata (x,y) non gode della proprietà $\mathcal{R}$. La $\mathcal{R}$ rappresenta dunque un legame tra certe coppie di elementi di S, e perciò si chiama una relazione binaria in S.

Una relazione binaria in S si chiama una relazione di equivalenza in S quando, come l'uguaglianza, soddisfa alle seguenti condizioni:

1) E' riflessiva, ossia risulta $xRx \ \forall x \in S$.

2) E' simmetrica, ossia tale che $xRy \Rightarrow yRx$

3) E' transitiva, ossia tale che xRy e $yRz \Rightarrow xRz$.

Se R è una relazione di equivalenza in S, e se x, y sono elementi di S, in luogo della scrittura xRy si adopera la seguente:

$$x \equiv y \ (R)$$

che più si avvicina alla scrittura $x = y$, e si legge "x equivalente y rispetto ad R", oppure "x equivalente y modulo R". Qualora la R sia stata fissata una volta per tutte, si può anche scrivere soltanto $x \equiv y$ e sottintendere il riferimento ad R.

Il giocatore di roulette che punta sul pari introduce in N la relazione di equivalenza R così definita:

« $x \equiv y \ (R)$ se e solo se x ed y sono entrambi pari o entrambi dispari »

e ripartisce l'insieme N in due classi: la classe N_1 dei numeri dispari, che sono a due a due equivalenti rispetto ad R, e la classe N_2 dei numeri pari, anch'essi a due a due equivalenti rispetto ad R. Gli insiemi N_1, N_2 si chiamano le classi di equivalenza rispetto ad R degli elementi di N, e costituiscono una partizione [1] di N. Al giocatore in questione conviene sostituire l'insieme N con l'insieme formato dalle due classi, cioè con l'insieme $\{N_1, N_2\}$ costituito da due elementi: tale insieme prende il nome di insieme quoziente di N rispetto alla relazione di equivalenza R. Ciascun numero dispari [pari] si chiama poi un rappresentante della classe N_1 [N_2], giacché la classe può essere "rappresentata" da uno qualunque dei suoi elementi.

Prima di precisare in generale le locuzioni introdotte, ci sembra opportuno considerare degli ulteriori esempi.

ESEMPIO 1. La relazione binaria R nell'insieme N definita come segue:

« xRy se e solo se x ed y hanno lo stesso numero di cifre »

è una relazione di equivalenza in N. Si ha ad esempio: $2 \equiv 5 \ (R)$, $31 \equiv 48 \ (R)$, $15 \not\equiv 8 \ (R)$.

(1) Il concetto di partizione di un insieme è noto al lettore. Tuttavia sarà richiamato in nota a pag. 6.

La classe di equivalenza del numero 1, diciamola A_1, e' costituita da tutti i numeri di una sola cifra, e naturalmente coincide con la classe di equivalenza del numero 2, del numero 3, ecc. Se per ogni $k \in N$ si considera un numero $a_k \in N$ di k cifre, la classe di equivalenza di a_k, diciamola A_k, e' costituita da tutti i numeri di k cifre, e coincide con la classe di equivalenza di un qualunque altro numero a'_k di k cifre.

Le classi di equivalenza sono allora rappresentate dai termini della successione $A_1, A_2, \dots, A_n, \dots$, che si denota anche con $(A_n)_{n \in N}$. L'insieme costituito dalle classi A_k, il quale insieme e' una partizione di N, e' l'insieme quoziente di N rispetto alla relazione di equivalenza $\mathcal{R}$.

I numeri di k cifre, elementi della classe A_k, sono i rappresentanti della classe.

ESEMPIO 2. La relazione binaria $\mathcal{R}$ definita nell'insieme R come segue:

« $x \mathcal{R} y$ se e solo se $\exists k \in Z : x - y = 2k\pi$ »

e' una relazione di equivalenza in R. Si ha ad esempio:

$$x \equiv x + 8\pi \ (\mathcal{R}) \ , \quad x \not\equiv x + 3\pi \ (\mathcal{R}).$$

Per ogni $x \in R$ la classe di equivalenza di x, diciamola A_x, e' costituita da tutti i numeri reali del tipo $x + 2k\pi$ con $k \in Z$. Naturalmente, se x ed y sono numeri reali distinti, non e' detto che la classe A_x sia diversa dalla classe A_y: se esiste $k \in Z$ tale che $x - y = 2k\pi$, cioe' se $x \equiv y \ (\mathcal{R})$, risulta $A_x = A_y$; in caso contrario, non solo si ha $A_x \neq A_y$, ma le classi A_x, A_y non hanno alcun elemento in comune.

Considerando la famiglia di classi $(A_x)_{x \in R}$, che questa volta non e' una successione, l'insieme $\mathcal{A}$ da esse costituito, cioe' l'insieme di tutte le possibili classi a due a due distinte (quindi anche disgiunte) e' l'insieme quoziente di R rispetto alla relazione di equivalenza $\mathcal{R}$. Ovviamente l'insieme $\mathcal{A}$ e' una partizione di R.

ESEMPIO 3. Fissato un intero positivo m, nell'insieme Z consideriamo la relazione binaria $\mathcal{R}_m$ definita come segue:

« $x \mathcal{R}_m y$ se e solo se $x-y$ è divisibile per m ».

La $\mathcal{R}_m$ è una relazione di equivalenza in Z (congruenza modulo m). Si ha ad esempio:

$$3x \equiv 6x \ (\mathcal{R}_3) \ , \quad 4x \equiv 12x^3 \ (\mathcal{R}_4) \ , \quad 5x \equiv 25x^2 \ (\mathcal{R}_5) \ ,$$

ma si preferisce scrivere:

$$3x \equiv 6x \ (\mathrm{mod}\, 3) \ , \quad 4x \equiv 12x^3 \ (\mathrm{mod}\, 4) \ , \quad 5x \equiv 25x^2 \ (\mathrm{mod}\, 5) \ .$$

Le classi di equivalenza rispetto ad $\mathcal{R}_m$ sono m insiemi $A_0, A_1, \ldots, A_{m-1}$ (che si chiamano classi dei resti modulo m): A_0 è l'insieme dei numeri interi divisibili per m, A_1 è l'insieme dei numeri interi che divisi per m danno resto 1, A_2 l'insieme dei numeri interi che divisi per m danno resto 2, ecc. L'insieme quoziente di Z rispetto alla relazione di equivalenza $\mathcal{R}_m$ è $\{A_0, A_1, \ldots, A_{m-1}\}$, ed è ovviamente una partizione di Z.

ESEMPIO 4. Fissato un intervallo $[a,b]$, denotiamo con $C^{(1)}$ l'insieme costituito da tutte le funzioni reali definite in $[a,b]$ e ivi dotate di derivata prima continua.

La relazione binaria $\mathcal{R}$ definita in $C^{(1)}$ come segue:

« $\varphi \mathcal{R} \psi$ se e solo se $\varphi' = \psi'$ »

è una relazione di equivalenza in $C^{(1)}$. Se φ è un'arbitraria funzione appartenente a $C^{(1)}$, la classe di equivalenza di φ è costituita da tutte le funzioni ψ tali che $\psi' = \varphi'$, cioè da tutte le funzioni ψ del tipo $\varphi + c$, con c costante. In altri termini, la classe di equivalenza di φ è l'integrale indefinito della funzione φ'.

Detto $C^{(0)}$ l'insieme delle funzioni continue nell'intervallo $[a,b]$, siccome per ogni $f \in C^{(0)}$ esistono funzioni $\varphi \in C^{(1)}$ tali che $\varphi' = f$ (ossia funzioni φ primitive di f), l'insieme quoziente di $C^{(1)}$ rispetto alla relazione di equivalenza $\mathcal{R}$ è l'insieme degli elementi della famiglia (cfr. pag. 19):

$$\left(\int f(x)\,dx \right)_{f \in C^{(0)}} .$$

Allo scopo di chiarire il ruolo dell'insieme quoziente $C^{(1)}/\mathcal{R}$, denotiamo con D la corrispondenza che ad ogni $\varphi \in C^{(1)}$ associa la derivata $\varphi' \in C^{(0)}$:

$$D : \varphi \in C^{(1)} \longrightarrow \varphi' \in C^{(0)} .$$

La D non è invertibile[1], giacché per ogni $f \in C^{(0)}$ l'equazione nell'incognita $\varphi \in C^{(1)}$:

$$D\varphi = f$$

ammette infinite soluzioni. Per ogni $\varphi \in C^{(1)}$ conveniamo allora di denotare con $\tilde{\varphi}$ la classe di equivalenza di φ rispetto ad $\mathcal{R}$ [ovviamente se $\psi = \varphi + c$ con c costante si ha $\tilde{\psi} = \tilde{\varphi}$], e successivamente conveniamo di chiamare "derivata della classe $\tilde{\varphi}$" la derivata di un qualunque rappresentante della classe stessa (tutti i rappresentanti hanno la stessa derivata).

Dopo ciò, se si denota con D^* la corrispondenza che ad ogni $\tilde{\varphi} \in C^{(1)}/\mathcal{R}$ associa la derivata di $\tilde{\varphi}$, per ogni $f \in C^{(0)}$ l'equazione nell'incognita $\tilde{\varphi} \in C^{(1)}/\mathcal{R}$:

$$D^* \tilde{\varphi} = f$$

ammette un'unica soluzione, che è l'integrale indefinito di f. Pertanto la funzione D^* è invertibile, e la sua inversa è la funzione che ad ogni $f \in C^{(0)}$ associa l'integrale indefinito di f.

Riassumendo: l'operazione di derivazione definita in $C^{(1)}$ non è invertibile; se però si considera l'insieme quoziente $C^{(1)}/\mathcal{R}$ e si definisce ivi tale operazione (nel modo naturale dianzi precisato), si ottiene la corrispondenza D^* che è invertibile.

* * *

Consideriamo ora un arbitrario insieme S, ed una relazione di equivalenza $\mathcal{R}$ in S.

Per ogni $x \in S$, l'insieme S_x costituito da tutti gli elementi di S equivalenti ad x rispetto ad $\mathcal{R}$:

$$S_x = \{ y \in S : y \equiv x \ (\mathcal{R}) \}$$

si chiama <u>la classe di equivalenza di x rispetto ad $\mathcal{R}$</u> (o <u>modulo $\mathcal{R}$</u>); ogni elemento di S_x si chiama un <u>rappresentante</u> della classe di equivalenza.

Evidentemente, se $x_1 \equiv x_2 \ (\mathcal{R})$, le classi di equivalenza di x_1 e di x_2, ossia

(1) Il concetto di funzione invertibile è noto al lettore. Tuttavia sarà richiamato a pag. 8.

gli insiemi S_{x_1} ed S_{x_2} coincidono: infatti per ogni $y \in S_{x_1}$ si ha $y \equiv x_1\ (R)$, ed essendo per ipotesi $x_1 \equiv x_2\ (R)$, per la transitività di R si ha $y \equiv x_2\ (R)$, ossia $y \in S_{x_2}$; ciò prova che $S_{x_1} \subseteq S_{x_2}$, e con lo stesso discorso si prova l'inclusione opposta, tenuto conto della simmetria di R [2) pag. 2].

L'insieme costituito dalle classi di equivalenza degli elementi di S, diciamolo Σ, è una partizione di S (1): essendo ovvio che l'unione di tali classi è uguale ad S (ogni elemento di S appartiene alla classe di equivalenza di se stesso), dimostriamo che tali classi sono a due a due disgiunte. A tale scopo siano S_{x_1} ed S_{x_2} due classi di equivalenza distinte, sicché $x_1 \not\equiv x_2\ (R)$: se esse avessero in comune un elemento y, si avrebbe $y \equiv x_1\ (R)$ e $y \equiv x_2\ (R)$, e quindi $x_1 \equiv x_2\ (R)$ contro l'ipotesi.

L'insieme Σ costituito dalle classi di equivalenza degli elementi di S si chiama <u>l'insieme quoziente di S rispetto ad R</u>, e si denota con $\frac{S}{R}$ o con S/R. Gli elementi di tale insieme formano, come dianzi dimostrato,

---

(1) Dato un sottoinsieme S_0 di un insieme S, e data una famiglia (cfr. pag. 19) di sottoinsiemi non vuoti di S:

$$(S_i)_{i \in I}$$

l'insieme degli elementi della famiglia si chiama una <u>partizione</u> di S_0 quando risulta:

$$\bigcup_{i \in I} S_i = S_0\ ,$$

ed inoltre due qualunque elementi distinti della famiglia sono disgiunti. Se non si suppone verificata quest'ultima condizione, e si in luogo della precedente uguaglianza si ha l'inclusione:

$$\bigcup_{i \in I} S_i \supseteq S_0,$$

si dice che l'insieme degli elementi della famiglia è un <u>ricoprimento</u> di S_0.

Gli elementi di una partizione di S_0 sono ovviamente sottoinsiemi di S_0, mentre gli elementi di un ricoprimento di S_0 sono sottoinsiemi di S che non necessariamente sono contenuti in S_0.

una partizione di S, che si può chiamare la partizione di S generata dalla relazione di equivalenza $\mathcal{R}$.

Vogliamo ora far vedere che, assegnata un'arbitraria partizione $\mathcal{P}$ dell'insieme S, esiste sempre una relazione di equivalenza $\mathcal{R}$ che genera tale partizione: sicché assegnando una partizione di S si viene automaticamente ad assegnare una relazione di equivalenza in S.

Assegnata invero una partizione $\mathcal{P}$ dell'insieme S, cioè un insieme di parti non vuote di S a due a due disgiunte ed aventi per unione S, denotiamo con $\mathcal{R}$ la relazione binaria in S definita come segue:

« $x \mathcal{R} y$ se esiste un elemento P della partizione $\mathcal{P}$ tale che $x \in P$ ed $y \in P$ ».

Ovviamente la $\mathcal{R}$ è riflessiva, simmetrica e transitiva, sicché $\mathcal{R}$ è una relazione di equivalenza in S. D'altra parte qualunque sia $x \in S$, per definizione di partizione esiste un unico $P \in \mathcal{P}$ al quale x appartiene, e la classe di equivalenza di x rispetto ad $\mathcal{R}$:

$$S_x = \{ y \in S : y \equiv x \ (\mathcal{R}) \}$$

ovviamente coincide con P. Dunque l'insieme quoziente $S/\mathcal{R}$, cioè l'insieme degli elementi della famiglia $(S_x)_{x \in S}$, coincide con la partizione $\mathcal{P}$: in altri termini, la partizione di S generata da $\mathcal{R}$ è proprio $\mathcal{P}$.

* * *

Per concludere, possiamo dire che una relazione di equivalenza $\mathcal{R}$ in S dà luogo nell'insieme quoziente $S/\mathcal{R}$ alla relazione di uguaglianza. Considerare l'insieme quoziente è come identificare tutti gli elementi di S ha loro equivalenti rispetto ad $\mathcal{R}$: poiché sul piano logico tale identificazione non è lecita, gli elementi equivalenti vengono per così dire "raggruppati" formando un nuovo ente, la classe di equivalenza; e le classi di equivalenza, che costituiscono una partizione di S, formano un nuovo insieme, l'insieme quoziente.

8

INSIEMI EQUIPOTENTI

Premesse.

Se S e T sono insiemi non vuoti, una funzione f definita in S ed a valori in T:

$$f : S \to T$$

si chiama anche una applicazione di S in T.

Con tale locuzione non si precisa se il codominio di f è una parte propria di T oppure coincide con T. Nei casi in cui occorre tale precisazione, per intendere che il codominio di f coincide con T si sostituisce nella precedente locuzione la preposizione "in" con "su": cioè si dice che f è un' applicazione di S su T (applicazione suriettiva).

Naturalmente non è escluso che esistano elementi distinti di S aventi lo stesso corrispondente tramite la legge f; in altri termini, può accadere che non sussista l'implicazione:

$$(1) \qquad x' \neq x'' \implies f(x') \neq f(x'').$$

Se vale la (1), si dice che f è una funzione invertibile, oppure che f è un'applicazione iniettiva, o un'applicazione biunivoca di S in T; se poi f è anche suriettiva, e se si vuole evidenziare tale circostanza, si dice che f è un' applicazione biunivoca di S su T.

Per una funzione invertibile $f : S \to T$ si definisce la funzione inversa, che si denota con f^{-1}: si tratta della funzione definita nel codominio di f, diciamolo $T_0 \subseteq T$, che ad ogni $y \in T_0$ fa corrispondere (associa, come suol dirsi) l'unica soluzione (elemento di S) dell'equazione $f(x) = y$ nell'incognita x.

Naturalmente la funzione f^{-1} è un' applicazione suriettiva, ed è biunivoca: ossia, è un' applicazione biunivoca di T_0 su S.

Può accadere che, dati due insiemi non vuoti S e T, non esista alcuna applicazione biunivoca di S in T, oppure, esistendo un' applicazione siffatta,

che non esista alcuna applicazione biunivoca di S su T. Ad esempio, se

$$S = \{1,2,3,4\} \quad , \quad T = \{1,2,3\}$$

ovviamente non esiste alcuna applicazione biunivoca di S in T; invece, se

$$S = \{1,2,3\} \quad , \quad T = \{1,2,3,4\},$$

esistono applicazioni biunivoche di S in T, ma nessuna di esse è suriettiva.

Considerazioni introduttive.

Dati due insiemi finiti, ci si può domandare se essi hanno lo stesso numero di elementi, e in caso contrario quale dei due ha più elementi dell'altro. Se i due insiemi sono infiniti, ovviamente la domanda è priva di senso: adesempio considerando l'insieme dei triangoli di un piano e l'insieme delle funzioni reali continue in un intervallo, non ha significato parlare del numero degli elementi dei due insiemi, e neppure domandarsi se il primo insieme ha più elementi del secondo.

La predetta domanda, in sostanza, induce ad effettuare un certo confronto tra due insiemi finiti. Orbene, se la domanda si formula opportunamente in termini diversi, tale confronto si può estendere agli insiemi infiniti; per comprendere come si perviene alla formulazione accennata, approfondiamo il confronto stesso nel caso di due insiemi finiti S e T.

Il modo più naturale di procedere è quello di contare gli elementi dei due insiemi, e poi confrontare i due numeri ottenuti, ma ovviamente questo procedimento non si può estendere al caso di due insiemi infiniti. Vi è un altro modo di procedere, che vogliamo illustrare prima con riferimento ad un esempio.

Supponiamo che S sia un insieme di studenti, e che T sia un insieme di libri. Immaginiamo allora di invitare gli studenti a prendere ciascuno un libro (uno solo), e che gli studenti procedano ad effettuare tale operazione.

Se restano alcuni studenti senza il libro, perché i libri si sono esauriti, ciò vuol dire che il numero degli studenti è maggiore di quello dei libri; se invece tutti hanno preso il libro e restano ancora dei libri, ciò vuol dire che il numero dei libri è maggiore di quello degli studenti; se infine tutti hanno preso il libro, e non vi sono più libri, ciò vuol dire che i due insiemi S e T hanno lo stesso numero di elementi.

Da notare, ed è questo il fatto importante, che il procedimento esposto consente di confrontare il numero degli elementi di S con quello degli elementi di T <u>senza determinare il numero degli elementi</u> di ciascuno dei due insiemi; in altri termini, con questo procedimento si può fare il confronto senza che si conosca <u>il numero</u> degli elementi dei due insiemi. Il procedimento è perciò suscettibile di estensione al caso di insiemi infiniti. Dobbiamo però, ancora supponendo che S e T siano finiti, esprimere il procedimento stesso indipendentemente dall'esempio considerato.

A tale scopo osserviamo che, invitando gli studenti a prendere ciascuno un libro, si cerca di stabilire una corrispondenza che ad ogni elemento di S (studente) associ un elemento di T (libro):

$$(2) \qquad f : S \to T,$$

e siccome non si deve consentire che due studenti prendano lo stesso libro da tenere in comune, la funzione f che si cerca di determinare deve soddisfare alla condizione:

$$x' \neq x'' \Longrightarrow f(x') \neq f(x''),$$

cioè deve essere biunivoca.

Se il numero degli studenti è minore o uguale a quello dei libri, e ciascuno sceglie un libro, si viene a determinare un'applicazione biunivoca f di S in T. In caso contrario, poiché restano studenti (elementi di S) senza libro, si viene a constatare che la corrispondenza cercata non esiste: volendo associare un libro anche agli studenti restanti, bisogna consentire che vi siano coppie di studenti con un libro in comune; cioè, esistono certamente delle applicazioni di S in T, ma non ne esiste una <u>biunivoca</u>. Però in tal caso (numero degli studenti maggiore di quello dei libri), si può considerare la corrispondenza che ad ogni libro (elemento di T) associa lo studente (elemento di S) che ha scelto quel libro, diciamola g:

(3) $$g : T \to S$$

e questa è biunivoca (non essendo consentito che due libri siano scelti dallo stesso studente). Perciò se il numero degli elementi di T è minore del numero degli elementi di S, non esiste alcuna applicazione biunivoca di S in T ma esiste un'applicazione biunivoca di T in S (peraltro l'esistenza di un'applicazione siffatta non esclude che S e T abbiano lo stesso numero di elementi).

Ovviamente, se i due insiemi hanno lo stesso numero di elementi, esistono sia un'applicazione biunivoca del tipo (2) sia un'applicazione biunivoca del tipo (3). Anzi in tal caso, essendo f invertibile e suriettiva, si può scegliere come funzione g l'inversa di f.

Le considerazioni svolte suggeriscono il modo di formulare la domanda posta all'inizio senza che intervenga <u>il numero</u> degli elementi dei due insiemi. Invece che domandarci se il numero degli elementi di S è minore del numero degli elementi di T, possiamo domandarci se esiste un'applicazione

biunivoca di S in T : in caso affermativo, possiamo dire che il numero degli elementi di S è ≤ del numero degli elementi di T senza aver "contato" gli elementi dei due insiemi. In questi termini perciò la domanda ha significato anche se riferita a due insiemi infiniti.

Naturalmente l'eventuale risposta affermativa si dovrà esprimere senza parlare del "numero" degli elementi ; in luogo della parola "numero" si adopera allora, come vedremo, la parola "potenza".

Definizioni.

Siano S, T insiemi non vuoti, finiti o infiniti. Se esiste un'applicazione biunivoca di S in T, si dice che la potenza di S è minore o uguale della potenza di T, oppure che la potenza di T è maggiore o uguale della potenza di S. Di conseguenza, se esiste un'applicazione biunivoca di T in S, si dirà che la potenza di T è minore o uguale della potenza di S, oppure che la potenza di S è maggiore o uguale della potenza di T.

Sorge spontanea la domanda: e se applicazioni biunivoche non ne esistono, né di S in T, né di T in S, come si fa il confronto? Orbene, tale circostanza non può verificarsi, in virtù del seguente teorema di Hartogs (equivalente al cosiddetto "assioma della scelta", di cui qui non ci occupiamo): qualunque siano gli insiemi non vuoti S e T, o esiste un'applicazione biunivoca di S in T, o esiste un'applicazione biunivoca di T in S.

E se si verificano simultaneamente l'una e l'altra di tali eventualità? In questa ipotesi, un teorema di Bernstein assicura che esiste un'applicazione biunivoca di S su T (quindi anche un'applicazione biunivoca di T su S).

Quando esiste un'applicazione biunivoca di S su T, si dice che gli insiemi S e T hanno la stessa potenza, oppure che S e T sono equipotenti.

Se la potenza di S è minore o uguale della potenza di T, e se inoltre S e T non sono equipotenti, allora si dice che la potenza di S è minore della potenza di T, oppure che la potenza di T è maggiore della potenza di S.

Le nozioni introdotte (che non si fondano sulle considerazioni introduttive cui ci siamo, per così dire, "ispirati") consentono di dare significato rigoroso alle locuzioni "insieme finito", "insieme infinito", "numero degli elementi" di un insieme finito. Orbene, un insieme non vuoto S si dice finito quando esiste $n \in N$ tale che S sia equipotente all'insieme $\{1, 2, \dots, n\}$. Se esiste un siffatto $n \in N$, ovviamente unico, esso si chiama il numero cardinale di S, oppure il numero degli elementi di S, e si denota con $\operatorname{card} S$.

Se S è un insieme finito, per ricercare il numero cardinale di S basta ovviamente associare ad un elemento di S il numero 1, ad un altro il numero 2, e così via fino a quando si esauriscono gli elementi di S: il numero n associato all'ultimo elemento è ovviamente $\operatorname{card} S$, in quanto la corrispondenza ottenuta è un'applicazione biunivoca di S sull'insieme $\{1, \dots, n\}$. È interessante notare che l'operazione del contare gli oggetti di un insieme finito è appunto la costruzione della predetta corrispondenza.

Se l'insieme non vuoto S non è finito, si dice che S è infinito: in tal caso, qualunque sia $n \in N$ la potenza di S è maggiore della potenza dell'insieme $\{1, \dots, n\}$. Ovviamente un insieme equipotente ad N è infinito; si dimostra che un insieme S è infinito se e solo se la potenza di S è maggiore o uguale della potenza di N.

Se S è un insieme finito, ovviamente nessuno dei sottoinsiemi propri di S ha lo stesso numero di elementi dell'insieme S, cioè non esiste alcun sottoinsieme proprio S_0 di S tale che $\text{card}\, S_0 = \text{card}\, S$. In altri termini, un insieme finito non può essere equipotente ad una sua parte propria.

Invece se S è un insieme infinito, esiste sempre una parte propria di S che è equipotente ad S. Si dimostra infatti che condizione necessaria e sufficiente affinché un insieme S sia infinito, è che S sia equipotente ad una sua parte propria.

Affinché ci si renda conto di tale risultato, è utile considerare il seguente esempio. Detto P il sottoinsieme proprio di N costituito dai numeri naturali pari, consideriamo la funzione $f : n \in N \to 2n \in P$; ovviamente se $n' \neq n''$ si ha $f(n') \neq f(n'')$, sicché f è biunivoca; inoltre il codominio di f, costituito da numeri pari, coincide con P, in quanto se m è un qualunque numero pari risulta $\frac{m}{2} \in N$ ed $f\left(\frac{m}{2}\right) = m$. Dunque esiste un'applicazione biunivoca di N su P, e perciò N e P sono equipotenti, pur essendo P una parte propria di N.
Naturalmente vi sono anche parti proprie di N non equipotenti ad N: basta considerare una qualunque parte finita di N.

La potenza di un insieme.

La locuzione introdotta per esprimere l'esistenza di un'applicazione biunivoca di un insieme in un altro, dal punto di vista logico non appare corretta nella forma in quanto vi interviene la parola "potenza", soggetto della frase, e non è stato precisato cosa debba intendersi per potenza di un insieme.

Per definire il termine "potenza", consideriamo un insieme I e denotiamo con $\mathcal{P}(I)$ l'insieme costituito da tutte le parti non vuote di I.

Introduciamo in $\mathcal{P}(I)$ la relazione binaria $\mathcal{R}$ definita come segue:
« gli insiemi S e T sono nella relazione $\mathcal{R}$ se e solo se essi sono equipotenti ».
Ovviamente la relazione $\mathcal{R}$ è riflessiva, simmetrica e transitiva: ossia, è una relazione di equivalenza nell'insieme $\mathcal{P}(I)$. Consideriamo allora l'insieme delle classi di equivalenza, rispetto ad $\mathcal{R}$, degli elementi di $\mathcal{P}(I)$, cioè l'insieme quoziente di $\mathcal{P}(I)$ rispetto ad $\mathcal{R}$; per ogni $S \in \mathcal{P}(I)$, denotiamo con $\widetilde{S}$ la classe di equivalenza di S, ossia la totalità delle parti di I equipotenti ad S.
Orbene, se S è un sottoinsieme di I, la classe di equivalenza $\widetilde{S}$ si chiama la potenza di S. Evidentemente, se S_1 è un sottoinsieme di I equipotente ad S, risulta $\widetilde{S}_1 = \widetilde{S}$, cioè la potenza di S_1 è uguale alla potenza di S.
Nell'insieme quoziente di $\mathcal{P}(I)$ rispetto ad $\mathcal{R}$, che si chiama l'insieme delle potenze delle parti di I, introduciamo la relazione binaria $\mathcal{M}$ definita come segue:
« $\widetilde{S}$ è nella relazione $\mathcal{M}$ con l'insieme $\widetilde{T}$ se e solo se la potenza di S è minore o uguale della potenza di T ».
La $\mathcal{M}$ è una relazione d'ordine nell'insieme delle potenze. Associando all'insieme delle potenze tale relazione d'ordine, esso diviene un insieme ordinato. Esiste l'elemento minimo, che è la potenza dei sottoinsiemi di I costituiti da un solo elemento, ed esiste pure l'elemento massimo, che è la potenza dei sottoinsiemi di I equipotenti ad I (se I è finito, l'unico sottoinsieme di I equipotente ad I è l'insieme I).
Da notare che si può parlare di potenza con riferimento ad un insieme qualunque, giacché qualunque insieme può essere considerato come sottoinsieme di un insieme I.

E' naturale domandarsi perché, allo scopo di definire la potenza di un insieme, si è fatto riferimento ai sottoinsiemi di un insieme I. Sarebbe stato più semplice introdurre la relazione di equivalenza $\mathcal{R}$, invece che in $\mathcal{P}(I)$, nell'insieme $\mathcal{T}$ costituito da tutti gli insiemi. Formalmente il discorso rimane inalterato se si sostituisce $\mathcal{P}(I)$ con $\mathcal{T}$, ma sul piano logico il discorso non avrebbe senso, perché non è lecito considerare l'insieme di tutti gli insiemi.

Per renderci conto di tale affermazione, ammettiamo che abbia senso l'insieme $\mathcal{T}$ costituito da tutti gli insiemi. Siccome anche $\mathcal{T}$ è un insieme, si ha che $\mathcal{T} \in \mathcal{T}$; siccome nessun elemento appartiene all'insieme vuoto, si ha che $\phi \notin \phi$.

Dunque la proprietà:

(4) $$X \notin X$$

è vera per $X = \phi$ ed è falsa per $X = \mathcal{T}$: ciò prova l'esistenza di insiemi X per cui la (4) è vera e di insiemi X per cui la (4) è falsa.

Orbene, è facile trovare un insieme S per cui la proprietà (4) non è né vera né falsa, il che è assurdo. Infatti, posto:

(5) $$S = \{ X \in \mathcal{T} : X \notin X \},$$

stante il significato di appartenenza all'insieme (5) non può verificarsi alcuna delle due eventualità $S \in S$, $S \notin S$ [l'eventualità $S \in S$ comporterebbe $S \notin S$, e l'eventualità $S \notin S$ comporterebbe $S \in S$]: pertanto per $X = S$ la (4) non è né vera né falsa.

Insiemi numerabili. Potenza del continuo.

La potenza di N si chiama la potenza del numerabile, e un insieme S equipotente ad N si dice un insieme numerabile.

Dunque un insieme S è numerabile se esiste un'applicazione biunivoca

f di N su S. Poiché le funzioni definite in N si chiamano successioni, dire che l'insieme S è numerabile significa dire che esiste una successione invertibile:

$$x_1, x_2, \ldots, x_n, \ldots \tag{6}$$

avente per codominio S. È chiara dunque l'origine della parola "numerabile": mediante la (6) si vengono a "numerare" gli elementi di S (x_1 = elemento n.1, x_2 = elemento n.2, ...).

Ovviamente un insieme infinito contiene sempre un insieme numerabile. Utilizziamo tale asserto per aggiungere un'osservazione sugli insiemi infiniti.

Se S è un insieme infinito, qualunque sia l'insieme finito T gli insiemi $S \cup T$ $S-T$ sono equipotenti ad S.

Per dimostrare che $S \cup T$ è equipotente ad S si può supporre, senza ledere la generalità, che gli elementi di T, diciamoli $y_1, \ldots, y_k$, non appartengano ad S. Scelto un insieme numerabile $M \subseteq S$, esistente poiché S è infinito, rappresentiamo gli elementi di M mediante la successione invertibile (6). Consideriamo quindi la funzione f definita nell'insieme $S \cup T$ come segue:

$$f(x) = x \quad \forall x \in S-M, \qquad f(x_1) = x_{k+1}, \; f(x_2) = x_{k+2}, \ldots, f(x_n) = x_{k+n}, \ldots$$
$$f(y_1) = x_1, \; f(y_2) = x_2, \ldots, f(y_k) = x_k$$

Evidentemente per $x' \neq x''$ si ha $f(x') \neq f(x'')$, e quindi f è biunivoca; inoltre il codominio di f, contenuto in S, coincide con S: invero un elemento x di S o appartiene ad $S-M$, ed allora è il corrispondente tramite f di se stesso, oppure appartiene ad M, e in tal caso, essendo uno dei termini della successione (6), diciamolo x_m, è il corrispondente tramite f dell'elemento x_{m-k} se $m > k$, dell'elemento y_m se $m \leq k$. In definitiva f è un'applicazione biunivoca di $S \cup T$ su S, e pertanto gli insiemi $S \cup T$ ed S sono equipotenti.

Per dimostrare che $S-T$ e' equipotente ad S si può supporre, senza ledere la generalità, che sia $T \subset S$. Essendo $S-T$ infinito in quanto S e' infinito e T e' finito, in base a quanto gia' dimostrato i due insiemi $S-T$, $(S-T) \cup T$ sono equipotenti: poiché il secondo di essi e' uguale ad S (in quanto $T \subset S$), si conclude che $S-T$ ed S sono equipotenti.

Si dimostra che:

1) Il prodotto cartesiano di due o piu' insiemi numerabili e' a sua volta un insieme numerabile.

2) Se gli insiemi della successione:

$$S_1, S_2, \ldots, S_n, \ldots$$

sono numerabili l'unione di tali insiemi e' numerabile.

L'insieme dei numeri razionali e' numerabile. Un sottoinsieme infinito di R^k costituito da punti razionali (ossia da punti aventi le coordinate razionali) e' numerabile.

Un insieme S che sia finito o numerabile si chiama un insieme discreto.

Abbiamo gia' detto che un insieme infinito ha potenza maggiore o uguale della potenza di N. E' importante osservare che effettivamente esistono insiemi aventi potenza maggiore di quella di N, ossia insiemi infiniti non numerabili: basta considerare l'intervallo $(0,1)$, che come e' noto non e' numerabile.

La potenza dell'intervallo $(0,1)$ si chiama la potenza del continuo. Siccome un qualunque intervallo non vuoto e non costituito da un solo punto e' equipotente all'intervallo $(0,1)$, ogni intervallo siffatto ha la potenza del continuo. Si dimostra che il prodotto cartesiano di k insiemi aventi la potenza del continuo ha anch'esso la potenza del continuo; di conseguenza R^k ha la potenza del continuo.

Ci si è posti la questione se esistano insiemi di potenza maggiore del numerabile e simultaneamente minore del continuo: si tratta di uno dei più importanti quesiti della Teoria degli insiemi. Cantor ha ipotizzato che la risposta sia negativa, e questa affermazione (generalmente accettata) viene chiamata ipotesi del continuo.

Esistono insiemi di potenza maggiore del continuo? La risposta, affermativa, è contenuta nel seguente teorema: qualunque sia l'insieme S, la potenza dell'insieme delle parti di S è maggiore della potenza di S.

* * *

Concludiamo le considerazioni svolte con una avvertenza.

Se si vogliono denotare con dei simboli gli oggetti di un insieme finito S, di solito si considera una lettera alfabetica affetta da indici. Adoperando ad es. la lettera x, se gli oggetti sono in numero di κ si possono denotare con

$$x_1, x_2, \ldots, x_\kappa .$$

Quando gli oggetti sono infiniti, si può essere istintivamente indotti a denotarli con una successione:

$$x_1, x_2, \ldots, x_n, \ldots .$$

Orbene, si tenga presente che se l'insieme S è infinito ma non è numerabile, non è lecito denotare i suoi elementi mediante una successione, perché non esiste alcuna successione invertibile avente per codominio S. Volendo rappresentare gli elementi di S occorrerà adoperare una notazione del tipo:

$$(7) \qquad (x_i)_{i \in I}$$

dove l'indice i varia in un insieme I (detto appunto insieme degli indici), che può essere l'intervallo $[0,1]$ o un insieme di potenza maggiore, naturalmente equipotente ad S.

Non precisando la potenza di I, la notazione (7) si può adoperare per denotare gli elementi di un qualunque insieme non vuoto, finito o infinito, diciamolo S. Si dice che la (7) è una famiglia di elementi di S, con insieme di indici I. La famiglia è una successione se I è numerabile, in particolare se $I = N$.

Capitolo I – Spazi topologici.

Introduzione.

In matematica la parola *spazio*, seguita sempre da un aggettivo o da un complemento di specificazione, sta ad indicare un insieme al quale è stato associato "qualcosa"; più precisamente sta ad indicare una coppia, di cui la prima coordinata è un insieme, e la seconda è qualcos'altro che viene richiamato dall'aggettivo, o dal complemento, che segue la parola *spazio*: si parla di spazio *numerico*, spazio *euclideo*, spazio *vettoriale*, spazio *mensurale*, spazio *di probabilità*, spazio *metrico*, spazio *normato*, ecc. Quello di cui dobbiamo occuparci in questo capitolo si chiama spazio *topologico*, e si ottiene associando ad un insieme qualunque una *topologia* ; si tratta di un concetto astratto, che costituisce il pilastro su cui poggia l'Analisi Funzionale: vogliamo farlo precedere da un discorso preliminare, cominciando col parlare di numeri.

* * *

Quando abbiamo a che fare con numeri che non sono interi , quasi sempre ricorriamo ad una *approssimazione*. Molte volte sopprimiamo le cifre decimali, e in questi casi adoperiamo la parola *arrotondamento* ; altre volte ne consideriamo solo alcune, per necessità dato che la natura umana non consente di scrivere infinite cifre decimali: ad esempio, quando dobbiamo calcolare l'area di un cerchio moltiplichiamo il quadrato del raggio per 3,14 , e spesso sorvoliamo sul fatto che il risultato non è l'area del cerchio, ma un valore approssimato dell'area. Se invece di moltiplicare per 3,14 scriviamo *moltiplicato* π , ci illudiamo di aver calcolato l'area con esattezza, mentre in sostanza l'abbiamo soltanto espressa simbolicamente.

Riflettiamo sul significato della parola *approssimazione*. Il numero 3,14 è un valore approssimato di π a meno di un centesimo: ciò vuol dire che la differenza dei due numeri , in valore assoluto, è minore di 1/100 . Se rappresentiamo su una retta i numeri reali e consideriamo il punto che rappresenta π, i valori approssimati per difetto o per eccesso di questo numero, a meno di 1/100 , sono quelli dell'intervallo aperto di estremi π-1/100, π+1/100 . Con il linguaggio dell'Analisi Matematica, si tratta dei punti di *un intorno* del punto π. Nella lingua inglese "un intorno" si dice "*a neighbourhood*", adoperando una parola che significa *vicinato* , ed esprime il fatto geometrico che si tratta di punti *vicini* al punto π. La vicinanza è quantificata dal numero 1/100, che è la semidimensione dell'intorno.

Dunque le approssimazioni di un numero x_0, o con linguaggio geometrico *di un punto* x_0, sono i punti di *un intorno* di x_0. Quando diciamo che una successione di numeri $x_1, x_2, \ldots, x_n, \ldots$ ha per limite il numero x_0, e scriviamo

$$\lim x_n = x_0 \quad , \tag{1}$$

sostanzialmente diciamo che x_0 può essere approssimato come vogliamo con i termini della successione; cioè che , se vogliamo un'approssimazione di x_0 *a meno di* ε*, cioè se fissiamo un intorno I di* x_0 *di semidimensione* ε, tutti i termini della successione da un certo indice in poi appartengono ad *I* e quindi realizzano l'approssimazione voluta.

Ricordiamo che il significato dell'uguaglianza (1) si esprime come segue:

(2) $$\forall\ I \textit{ intorno di } x_0,\ \exists\ \nu \in \mathbf{N}:\quad n > \nu \Rightarrow x_n \in I\ .$$

Dunque *approssimare* un numero x_0 significa scegliere una successione di numeri che ha per limite x_0.

Soffermiamoci ora sul significato di *approssimazione* con riferimento ad una funzione : ne abbiamo bisogno in diverse occasioni, perché spesso la funzione che ci interessa non si può esprimere, come suol dirsi, *in forma chiusa* , cioè non è *dotata di espressione elementare* . La parola *approssimazione*, riferita ad una funzione, contrariamente a quanto accade per un numero non ha un unico significato, ma ha significati diversi che dipendono dalla questione che si sta studiando. E' opportuno illustrare questo aspetto con un esempio.

Consideriamo l'equazione differenziale nell'incognita $x=x(t)$:

$$x'' + a(t)x' + b(t)x = f(t) \quad , \quad t \in [0,1]$$

con le condizioni iniziali $x(0) = x'(0) = 0$. Si tratta di ricercare una funzione $x_0(t)$, derivabile due volte nell'intervallo [0,1], tale da risultare:

(3) $$x_0''(t) + a(t)x_0'(t) + b(t)x_0(t) = f(t) \quad \forall\ t \in [0,1] \quad , \qquad x_0(0) = x_0'(0) = 0\,.$$

Immaginiamo di approssimare il termine noto f mediante una successione $f_1, f_2, \ldots, f_n, \ldots$, nel senso che per ogni fissato t si ha $\lim f_n(t) = f(t)$ (convergenza *puntuale* della successione di funzioni); immaginiamo poi di saper risolvere il problema solo quando nell'equazione si sostituisce f con f_n . Detta allora x_n la soluzione trovata , per ogni fissato t si avrà :

(4) $$x_n''(t) + a(t)x_n'(t) + b(t)x_n(t) = f_n(t) \quad , \qquad x_n(0) = x_n'(0) = 0 \quad .$$

Se per ogni fissato t risulta:

(5) $$\lim_{n\to\infty} x_n(t) = x_0(t)\ ,$$

non è detto che risulti anche

(6) $$\lim_{n\to\infty} x_n'(t) = x_0'(t) \quad , \qquad \lim_{n\to\infty} x_n''(t) = x_0''(t)$$

e quindi non è detto che dalle (4) conseguano le (3). Dunque se la successione di funzioni $x_1, x_2, \ldots, x_n, \ldots$ approssima x_0 nello stesso senso in cui $f_1, f_2, \ldots, f_n, \ldots$ approssima f , non è detto che la funzione x_0 sia soluzione del nostro problema. E' chiaro allora che, con riferimento alla soluzione del problema considerato, il significato della parola *approssimazione* deve coinvolgere non solo la funzione (come accade con riferimento al termine noto), ma anche le sue derivate prima e seconda : precisamente, dire che la successione $x_1, x_2, \ldots, x_n, \ldots$ approssima la soluzione del problema deve

significare non solo che si verifica la (5), ma anche che sussistono le (6). Prima di proseguire il nostro discorso, rileviamo che nell'esempio in questione è utile considerare l'insieme S costituito da tutte le funzioni definite nell'intervallo [0,1] ed ivi derivabili due volte : tale insieme viene chiamato *l'insieme ambiente* del problema, e si suol dire che il problema *è ambientato* in tale insieme.

Se poi consideriamo un' equazione differenziale di ordine 3, 4,… cambia *l'insieme ambiente* e la parola *approssimazione* dovrà avere un significato diverso, perché dovrà coinvolgere anche la derivata terza, la derivata quarta, … Pensando ad una ulteriore questione, può darsi che l'insieme ambiente non coinvolga le derivate: se approssimando i dati riusciamo a costruire una certa successione $x_1,x_2,\ldots,x_n,\ldots$, potrebbe non avere alcun interesse che per ogni fissato t la funzione x_n e le sue derivate abbiano rispettivamente per limite $x_0(t)$ e le sue derivate calcolate in t, ma invece abbia interesse approssimare la funzione $x(t)$ in media , oppure in media quadratica, cioè che per la successione x_n risulti:

$$(7) \qquad \lim \int_0^1 \left|x_n(t)-x_0(t)\right| dt = 0 \qquad \text{oppure} \qquad \lim \int_0^1 \left|x_n(t)-x_0(t)\right|^2 dt = 0 \quad .$$

Allora nell'insieme ambiente la parola *approssimazione* dovrà significare che sussiste la prima delle (7), oppure la seconda.

Da quanto si è detto si comprende che con riferimento ad una funzione, il significato della parola *approssimazione* non può essere sempre lo stesso : detto S l'insieme in cui è ambientato il problema, il significato dipende dall'insieme S e dal problema che si vuole studiare.

* * *

Molte volte abbiamo bisogno di approssimare anche altri oggetti : nella ricerca scientifica, ed anche nella pratica ad un certo livello, abbiamo bisogno di approssimare matrici, o vettori, o coppie di vettori, o coppie formate da un numero e da una funzione, oppure da una funzione e da una matrice, oppure terne di oggetti di questo tipo, e così via . Perciò abbiamo bisogno di parlare di *limite di una successione* anche se i suoi termini, ed il limite , non sono dei numeri ma oggetti di varia natura , elementi di un certo insieme S.

Allora riflettiamo sulla proposizione (2) , immaginando che x_0 ed x_n siano, invece che numeri, elementi dell' insieme S in cui è ambientato il nostro problema, e che I sia un sottoinsieme di S : immaginiamo di aver constatato quali oggetti ci interessa considerare vicini ad x_0, cioè quali oggetti debbano costituire nell'insieme ambiente *gli intorni* di x_0 . Dopo aver deciso quali sottoinsiemi di S ci interessa chiamare *intorni* di x_0 , la proposizione (2), ossia il limite di una successione, conserva il suo significato rigoroso: gli oggetti di un fissato intorno saranno *le approssimazioni* dell'oggetto x_0.

In definitiva: se alla parola *approssimazione* vogliamo dare un significato preciso con riferimento ad un elemento x_0 di un arbitrario insieme S , come abbiamo fatto con riferimento ad un numero, dobbiamo poter parlare di limite, oltre che per una successione di numeri, anche per una successione di elementi di S. Per raggiungere questo obiettivo, dobbiamo scegliere opportunamente , in base agli scopi da perseguire, i sottoinsiemi di S da chiamarsi *gli intorni* di un dato elemento x_0 di S.

Ipotizzando di aver fatto la scelta, dal significato della parola *intorno* , scelto per la questione che ci interessa, scaturiranno , oltre che la nozione di limite di una successione di elementi di *S* , tante altre nozioni, analogamente a quanto si è fatto nei corsi di Analisi Matematica (punto di accumulazione per un insieme, frontiera di un insieme, insieme chiuso, insieme aperto, ecc).

Peraltro, invece di partire dalla scelta degli intorni, per lo sviluppo della teoria è risultato conveniente supporre di avere scelto, in base al problema da affrontare, prima i sottoinsiemi di *S* da chiamarsi *insiemi aperti* , e poi dalla scelta ipotizzata [si badi bene: non abbiamo detto *effettuata*] far discendere le altre nozioni, compresa quella di *intorno* di un elemento di *S*. Notiamo che invece nell'Analisi 2 , dove *S* è l'insieme dei punti del piano, si introducono prima gli intorni e poi gli insiemi aperti , che sono quelli a cui i punti della frontiera non appartengono.

Naturalmente la classe $\mathcal{T}$ dei sottoinsiemi di *S* che si chiameranno *insiemi aperti,* deve essere assoggettata a certe condizioni: se non c'è alcuna ipotesi, non si potrà dimostrare nulla e non può nascere alcuna teoria.

1. Generalità.

Siano *S* un insieme e $\mathcal{P}(S)$ l'insieme costituito da tutte le parti di *S*.

Una classe $\mathcal{T}$ di sottoinsiemi di *S,* cioè una parte di $\mathcal{P}(S)$, si chiama una **topologia per** *S* quando soddisfa alle condizioni seguenti:

1. $\emptyset \in \mathcal{T}$ e $S \in \mathcal{T}$.
2. *L'intersezione di due elementi di* $\mathcal{T}$ *, e quindi di un numero finito di elementi di* $\mathcal{T}$ *, è a sua volta un elemento di* $\mathcal{T}$.
3. *Qualunque sia la famiglia* $(A_i)_{i\in I}$ *di elementi di* $\mathcal{T}$ *, l'insieme* $\bigcup_{i\in I} A_i$ *è a sua volta un elemento di* $\mathcal{T}$.

E' importante che il lettore si soffermi sul simbolo di "famiglia" che interviene nella condizione 3. L'insieme degli indici può essere ad esempio uno dei seguenti:

$$I = \{1, 2, \ldots, n\}\ , \quad I = N \quad , \quad I = R\ .$$

La "famiglia" nel primo caso è una ennupla di insiemi (coppia per n=2, terna per n=3), nel secondo è una successione ; nel terzo , se $A_i \neq A_j$ per $i \neq j$, la totalità degli insiemi della famiglia ha la potenza del continuo.

Il lettore osservi che anche nel secondo e terzo caso gli insiemi della famiglia possono essere in numero finito (può ad es. accadere che siano tutti uguali!) : naturalmente è un'eventualità che non ha alcun interesse nella 3.

Esempi di topologie per *S* sono l'insieme $\mathcal{T} = \{\emptyset, S\}$, che si chiama la **topologia banale per** *S* e l'insieme $\mathcal{T} = \mathcal{P}(S)$, che si chiama la **topologia discreta per** *S* .

Se $\mathcal{T}$ è una topologia per *S,* la coppia $(S,\mathcal{T})$ si chiama **spazio topologico**, del quale l'insieme *S* si chiama **il sostegno**. Gli elementi della classe $\mathcal{T}$ si chiamano **insiemi aperti** (o senz'altro **aperti**) dello spazio topologico $(S,\mathcal{T})$; gli elementi del sostegno *S* si chiamano **punti** dello spazio, oppure **punti di** *S*.

La terminologia che andiamo ad introdurre si intenderà riferita allo spazio topologico $(S,\mathcal{T})$.

Qualunque sia $X \subseteq S$, l'insieme *S-X* (*complementare di X*, o *complemento di X rispetto ad S*) sarà denotato con $-X$.

Si definisce **chiuso** ogni sottoinsieme C di S tale che $-C$ sia aperto, sicchè qualunque sia l'aperto A il suo complementare $-A$ è un insieme chiuso.

La classe degli insiemi chiusi, diciamola $\mathcal{C}_c$, gode delle seguenti proprietà:

1′. $\emptyset \in \mathcal{C}_c$ e $S \in \mathcal{C}_c$.

2′. *Qualunque sia la famiglia* $(C_i)_{i\in I}$ *di elementi di* $\mathcal{C}_c$ *, l'insieme* $\bigcap_{i\in I} C_i$ *è a sua volta un elemento di* $\mathcal{C}_c$.

3′. *L'unione di due elementi di* $\mathcal{C}_c$ *, e quindi di un numero finito di elementi di* $\mathcal{C}_c$ *, è a sua volta un elemento di* $\mathcal{C}_c$.

Tali proprietà si dimostrano facilmente in base alle cosiddette *leggi di De Morgan*([1]). Il lettore confronti le condizioni 2,3 imposte alla classe $\mathcal{C}$, con le proprietà 2′,3′ della classe $\mathcal{C}_c$: quella che nell'una si riferisce a due insiemi [ad una famiglia di insiemi], nell'altra si riferisce ad una famiglia di insiemi [a due insiemi].

Osserviamo che:

Qualunque siano A aperto e C chiuso, risulta A-C aperto e C-A chiuso.

DIM. Essendo $A\text{-}C = A \cap (-C)$, l'insieme $A\text{-}C$ è aperto perché intersezione di due aperti . Analogamente l'insieme $C\text{-}A$ è chiuso perché intersezione dell' insieme chiuso C e dell'insieme chiuso $-A$.

---

* * *

Consideriamo ora un punto x_0 di S e un sottoinsieme I di S.

Si dice che I è **un intorno** del punto x_0, quando I contiene un aperto al quale appartenga x_0 . Di conseguenza ogni insieme che contiene un intorno di x_0 è a sua volta un intorno di x_0 .

La nozione di aperto e quella di intorno sono collegate dal teorema che segue.

(1.1). *Un insieme è aperto se e solo se è un intorno di ogni suo punto.*

DIM. Sia X un insieme aperto, e sia x_0 un qualunque punto di X . Allora X contiene un aperto (cioè X stesso) al quale x_0 appartiene, e perciò X è un intorno di x_0 .

Viceversa, supponiamo che l'insieme X sia un intorno di ogni suo punto: allora a ciascun $x \in X$ si può associare un aperto $A_x \subseteq X$ al quale x appartenga. L'unione di tutti gli insiemi A_x , diciamola Y, è un sottoinsieme di X , ed è ovvio che $Y = X$ in quanto ogni $x \in X$ appartiene al corrispondente A_x e quindi ad Y . Dunque X è un' unione di aperti, e pertanto è un aperto.

([1]) Con tale locuzione si intendono le ovvie uguaglianze:

$$-\bigcup_{i\in I} X_i = \bigcap_{i\in I}(-X_i) \qquad , \qquad -\bigcap_{i\in I} X_i = \bigcup_{i\in I}(-X_i) \quad ,$$

che brevemente si possono esprimere dicendo che *il complementare dell'unione è uguale all'intersezione dei complementari, e il complementare dell'intersezione è uguale all'unione dei complementari*.

Siano x_0 un punto di S ed X un sottoinsieme di S.

Si dice che il punto x_0 è **interno ad** X quando esiste un intorno di x_0 contenuto in X .

L'insieme costituito dai punti interni ad X si chiama **l'interno di** X e si denota con

$$\overset{o}{X} \quad .$$

Si dice che il punto x_0 è **aderente ad** X quando a ciascun intorno di x_0 appartiene almeno un punto di X .

L'insieme costituito dai punti aderenti ad X si chiama **la chiusura** o **l'aderenza d**i X e si denota con

$$\overline{X} \quad .$$

Evidentemente si ha:

$$\overset{o}{X} \subseteq X \subseteq \overline{X} \quad . \tag{1}$$

Una caratterizzazione del'interno e della chiusura di un insieme è fornita dal teorema che segue.

(1.2). *L'interno di un insieme X è un aperto, ed è il più grande (rispetto all'inclusione) degli insiemi aperti contenuti in X . La chiusura di un insieme X è un chiuso, ed è il più piccolo (rispetto all'inclusione) degli insiemi chiusi contenenti X.*

Dette $\mathscr{A}$ la classe degli aperti contenuti in X e $\mathscr{C}$ la classe dei chiusi contenenti X , le precedenti asserzioni si deducono dalle uguaglianze:

$$\overset{O}{X} = \bigcup_{A \in \mathscr{A}} A \quad , \quad \overline{X} = \bigcap_{C \in \mathscr{C}} C \tag{2}$$

DIM. Ad ogni punto x interno ad X si può associare un intorno I_x di x contenuto in X, e I_x contiene un aperto al quale x appartenga : dunque x appartiene ad un aperto contenuto in X, ossia appartiene al secondo membro della prima delle (2). Pertanto, con riferimento alla prima delle (2), l'insieme a primo membro è incluso in quello a secondo membro. Per dimostrare l'inclusione opposta consideriamo un qualunque punto x_0 dell'insieme a secondo membro. Siccome x_0 appartiene ad un aperto A contenuto in X, l'insieme A è un intorno di x_0 contenuto in X, sicchè x_0 è interno ad X, ossia appartiene all'insieme a primo membro . La prima delle (2) è così acquisita.

La seconda delle (2) si stabilisce facilmente in base alle leggi di De Morgan [cfr. nota ([1]) pag.5], dopo aver osservato che il complementare della chiusura di X coincide con l'interno del complementare di X.

E' facile altresì dedurre dalle uguaglianze (2) la prima parte del teorema, tenendo presente la condizione 3 di pag.4 .

Il teorema ora acquisito consente di aggiungere un' importante precisazione alle due inclusioni che figurano nella (1) ; questa asserisce, per così dire, che qualunque insieme " è compreso " tra l'interno e la chiusura di esso. Orbene, sussiste il teorema:

(1.3). *Un insieme X è aperto se e solo se coincide con il suo interno, è chiuso se e solo se coincide con la sua chiusura.*

DIM. Se X è aperto, il più grande aperto contenuto in X è X stesso : ne segue, per (1.2), che X coincide con il suo interno. Viceversa, se X coincide con il suo interno, per (1.2) X è aperto. Per la seconda parte del teorema il ragionamento è del tutto analogo.

Siano ancora x_0 un punto di S ed X un sottoinsieme di S.

Si dice che x_0 è **un punto di accumulazione per l'insieme** X quando a ciascun intorno di x_0 appartiene almeno un punto di X che sia diverso da x_0 .

L'insieme costituito dai punti di accumulazione per X si chiama **il derivato d**i X e sarà denotato con

$$Dr(X) \ .$$

Può accadere che un punto di accumulazione per X non appartenga ad X, e che un punto di X non sia di accumulazione per X . Un punto x_0 che appartiene ad X e non è di accumulazione per X si chiama **punto isolato dell'insieme** X.

(1.4). *Qualunque sia il sottoinsieme X di S, si ha:*

$$\overline{X} = X \cup Dr(X) \ .$$

DIM. Ogni punto di accumulazione per X è anche punto aderente ad X, sicchè risulta $Dr(X) \subseteq \overline{X}$, e d'altra parte per la (1) risulta $X \subseteq \overline{X}$: pertanto $X \cup Dr(X) \subseteq \overline{X}$.

Per dimostrare l'inclusione opposta consideriamo un qualunque punto $x_0 \in \overline{X}$.

Se x_0 appartiene ad X, appartiene anche all'insieme $X \cup Dr(X)$; se x_0 non appartiene ad X, essendo punto aderente ad X deve accadere che a ciascun intorno di x_0 appartiene un punto di X, il quale non può essere x_0 , sicchè x_0 è di accumulazione per X e perciò appartiene all'insieme $X \cup Dr(X)$. Dunque è vera la predetta inclusione opposta, e con ciò l'asserto è provato.

(1.5). *Un insieme è chiuso se e solo se contiene il suo derivato.*

DIM. Tenendo conto delle (1.3),(1.4) si ha:

$$X \text{ chiuso} \Leftrightarrow X = \overline{X} \Leftrightarrow X = X \cup Dr(X) \Leftrightarrow X \supseteq Dr(X) \quad .$$

Si dice che x_0 è un **punto frontiera per l'insieme** X quando x_0 è aderente sia ad X che al complementare $-X$.

L'insieme costituito dai punti frontiera per X si chiama **la frontiera** di X e si denota con

$$\partial X \ .$$

Dalla definizione stessa di punto frontiera segue che:

$$(3) \qquad \partial X = \partial(-X) \quad , \quad \partial X = \overline{X} \cap \overline{(-X)}$$

e per la seconda delle (3) *la frontiera di un insieme, essendo intersezione di due chiusi, è un insieme chiuso.*

Nelle uguaglianze che seguono ciascuna delle nozioni di chiusura e di interno viene collegata con la nozione di frontiera .

(1.6). *Qualunque sia X sottoinsieme di S risulta*:

(4) $$\overline{X} = X \cup \partial X \quad , \quad \overset{o}{X} = X - \partial X \ .$$

DIM. Proviamo dapprima l'inclusione $\overline{X} \subseteq X \cup \partial X$, e perciò consideriamo un punto $x_0 \in \overline{X}$.

Se appartiene ad X, il punto x_0 appartiene anche ad $X \cup \partial X$; se non appartiene ad X, sicchè appartiene a $-X$, il punto x_0 è aderente sia ad X che a $-X$, quindi $x_0 \in \partial X$. In ogni caso $x_0 \in X \cup \partial X$ e con ciò l'inclusione è acquisita. Per dimostrare la prima delle (4) basta allora provare l'inclusione opposta, ossia $X \cup \partial X \subseteq \overline{X}$: ma questa consegue dalle (1), (3) [la seconda delle (3) implica $\partial X \subseteq \overline{X}$].

Quanto alla seconda delle (4), basta osservare che un punto interno ad X è un punto di X e non è di frontiera per X [quindi $\overset{o}{X} \subseteq X - \partial X$], e che un punto di X, se non è di frontiera, è necessariamente interno ad X [quindi $X - \partial X \subseteq \overset{o}{X}$].

Dalla (1.6) si deduce che:

(1.7). *Qualunque sia* $X \subseteq S$ *, si ha*:

(5) $$X \text{ chiuso} \Leftrightarrow \partial X \subseteq X \quad , \quad X \text{ aperto} \Leftrightarrow \partial X \subseteq -X \quad .$$

Di conseguenza :

(6) $$X \text{ simultaneamente aperto e chiuso} \Leftrightarrow \partial X = \varnothing \ .$$

X *né aperto né chiuso* $\Leftrightarrow$ *gli insiemi* $X \cap \partial X$, $(-X) \cap \partial X$ *sono entrambi non vuoti.*

DIM. Se $\partial X \subseteq X$, dalla prima delle (4) consegue $\overline{X} \subseteq X$ e quindi, essendo $X \subseteq \overline{X}$ per la (1), si ha $X = \overline{X}$ ossia X chiuso [per (1.3)] .Viceversa si ha X *chiuso* $\Rightarrow X = \overline{X} \supseteq \partial X \Rightarrow X \supseteq \partial X$ e con ciò la prima delle (5) è dimostrata. Da essa segue subito la seconda delle (5), tenuto conto della prima delle (3). Infatti:

$$X \text{ aperto} \Leftrightarrow -X \text{ chiuso} \Leftrightarrow \partial(-X) \subseteq -X \Leftrightarrow \partial X \subseteq -X \ .$$

Proviamo ora la (6). Se X è aperto e chiuso, per le (5) si ha :

$$\left[\partial X \subseteq X \cap (-X) = \varnothing\right] \Rightarrow \partial X = \varnothing \ .$$

Viceversa se ∂X è vuoto, è contenuto in X e in $-X$, e quindi per le (5) X è simultaneamente aperto e chiuso**.**

L'ultima affermazione è ovvia conseguenza delle (5).

Quando non occorre menzionare il sostegno e la topologia dello spazio topologico $(S, \mathcal{T})$, conviene indicare quest'ultimo, come noi faremo spesso nel seguito, scrivendo solo la lettera S . Ciò non dà luogo ad equivoci : ad esempio è chiaro che le parole punto, sottoinsieme, ecc., sono riferite al sostegno, e per un eventuale riferimento alla topologia si può dire "*la topologia di S* ".

In particolare si usa dire *lo spazio topologico* $\mathbf{R}^n$ (anche per $n = 1$) intendendo la coppia $(\mathbf{R}^n, \mathcal{T})$, dove $\mathcal{T}$ è la cosiddetta *topologia usuale per* $\mathbf{R}^n$ adoperata nell' Analisi Matematica : nei relativi corsi essa viene introdotta implicitamente, ed è la classe dei sottoinsiemi di $\mathbf{R}^n$ ciascuno dei quali si può esprimere come unione di una famiglia di intervalli aperti di $\mathbf{R}^n$ [si può dimostrare che $\mathcal{T}$ verifica le condizioni 1,2,3 di pag.4 con $S = \mathbf{R}^n$].

OSSERVAZIONE.

Gli insiemi $\emptyset$, S sono simultaneamente aperti e chiusi, e sono privi di punti frontiera. Se $S = \mathbf{R}^n$ non esistono altri insiemi simultaneamente aperti e chiusi, o equivalentemente [per la (6)] altri insiemi la cui frontiera sia l'insieme vuoto. E' naturale domandarsi se ciò accade in tutti gli spazi topologici: orbene, come vedremo nel n.3, la risposta è negativa.

2. Sottospazi. Spazi di Hausdorff.

Sia $(S, \mathcal{T})$ uno spazio topologico, e sia S_0 un sottoinsieme proprio e non vuoto di S .

Un insieme $X \subseteq S$ si dice **aperto su** S_0 , o **aperto in** S_0 [**chiuso su** S_0 , o **chiuso in** S_0] quando si può esprimere come intersezione di S_0 con un aperto [con un chiuso] ; ossia quando esiste un aperto A [un chiuso C] tale che :

$$X = S_0 \cap A \qquad [\ X = S_0 \cap C\ ] .$$

Analogamente, detto x_0 un punto di S_0 , un insieme I_0 si chiama **intorno di** x_0 **su** S_0 , oppure **intorno di** x_0 **in** S_0 , quando si può esprimere come intersezione di S_0 con un intorno di x_0 ; ossia quando esiste intorno I di x_0 tale che:

$$I_0 = S_0 \cap I .$$

Ovviamente un insieme che sia aperto su S_0 , o chiuso su S_0 , o intorno su S_0 di un punto di S_0 , è contenuto in S_0 .

Le locuzioni dianzi introdotte derivano dal fatto che *la totalità degli insiemi aperti su* S_0 *, diciamola* $\mathcal{T}_0$ *, è una topologia per* S_0 .

DIM. Si tratta di verificare che sono soddisfatte le condizioni 1,2,3 di pag.4.

1). $\emptyset \in \mathcal{T}_0$: infatti $\emptyset = S_0 \cap \emptyset$ e $\emptyset$ aperto ;
$S_0 \in \mathcal{T}_0$: infatti $S_0 = S_0 \cap S$ e S aperto.

2). Siano B_1 , B_2 elementi di $\mathcal{T}_0$, sicchè esistono due aperti A_1, A_2 per i quali si ha $B_1 = S_0 \cap A_1$, $B_2 = S_0 \cap A_2$. Ne consegue:
$B_1 \cap B_2 = (S_0 \cap A_1) \cap (S_0 \cap A_2) = S_0 \cap (A_1 \cap A_2)$ e $A_1 \cap A_2$ *aperto* ,
e pertanto $B_1 \cap B_2$ è elemento di $\mathcal{T}_0$.

3). Sia $(B_i)_{i \in I}$ una famiglia di elementi di $\mathcal{T}_0$, sicchè per ogni $i \in I$ esiste un aperto A_i per il quale si ha $B_i = S_0 \cap A_i$. Ne consegue:
$$\bigcup_{i \in I} B_i = \bigcup_{i \in I} (S_0 \cap A_i) = S_0 \cap \bigcup_{i \in I} A_i \text{ e } \bigcup_{i \in I} A_i \text{ } aperto ,$$
e pertanto $\bigcup_{i \in I} B_i$ è elemento di $\mathcal{T}_0$.

Orbene, la topologia $\mathcal{T}_0$ si dice **indotta su** S_0 **dalla topologia** $\mathcal{T}$, e lo spazio topologico $(S_0 , \mathcal{T}_0)$ si chiama **sottospazio di** $(S, \mathcal{T})$ **di sostegno** S_0 .

Per definizione, gli insiemi aperti del sottospazio $(S_0 , \mathcal{T}_0)$ sono gli insiemi aperti su S_0 . Dimostriamo che anche *gli insiemi chiusi del sottospazio* $(S_0 , \mathcal{T}_0)$ *sono gli insiemi chiusi su* S_0 , e che *gli intorni di un punto* $x_0 \in S_0$ *nel sottospazio* $(S_0 , \mathcal{T}_0)$ *sono gli intorni di* x_0 *su* S_0 .

DIM. Si ha:

$$X \text{ chiuso nel sottospazio } (S_0 , \mathcal{T}_0) \Leftrightarrow S_0\text{-}X \text{ aperto nel sottospazio } (S_0 , \mathcal{T}_0) \Leftrightarrow$$

$$S_0\text{-}X \text{ aperto su } S_0 \Leftrightarrow \text{esiste un aperto } A \text{ tale che } S_0\text{-}X = S_0 \cap A.$$

L'ultima uguaglianza, passando ai complementari rispetto ad S_0 , equivale alla seguente:

$$X = S_0 - (S_0 \cap A) = S_0 \cap (-A) .$$

Siccome $-A$ è chiuso , ciò significa che X è chiuso su S_0 , e la prima asserzione è dimostrata.

Per dimostrare la seconda asserzione, consideriamo un punto $x_0 \in S_0$ e un intorno I_0 di x_0 nel sottospazio $(S_0 , \mathcal{T}_0)$. Per la definizione di intorno, l'insieme I_0 contiene un insieme B, al quale x_0 appartiene, aperto nel sottospazio ossia del tipo $B = S_0 \cap A$ con A aperto. Notiamo che di conseguenza il punto x_0 appartiene ad A.

Tenendo presente che $S_0 \cap I_0 = I_0$ [in quanto $I_0 \subseteq S_0$] e che $(S_0 \cap A) \cup I_0 = I_0$ [in quanto $S_0 \cap A \subseteq I_0$] , si ha :

$$S_0 \cap (A \cup I_0) = (S_0 \cap A) \cup (S_0 \cap I_0) = (S_0 \cap A) \cup I_0 = I_0$$

e quindi :

(1) $$I_0 = S_0 \cap (A \cup I_0) .$$

Siccome $A \cup I_0$ è un intorno di x_0 [dato che contiene l'aperto A al quale, come dianzi evidenziato, x_0 appartiene] , la (1) esprime I_0 come intersezione di S_0 con un intorno di x_0 , e perciò I_0 è un intorno di x_0 su S_0 .

Viceversa, supponiamo che I_0 sia un intorno di I_0 su S_0, cioè che esista un intorno I di x_0 tale che $I_0 = S_0 \cap I$. Detto A un aperto contenuto in I al quale appartenga x_0 , si ha:

$$I_0 = S_0 \cap I \supseteq S_0 \cap A ,$$

e quindi I_0 contiene un aperto su S_0 al quale x_0 appartiene. Pertanto I_0 è un intorno di x_0 nel sottospazio $(S_0 , \mathcal{T}_0)$.

OSSERVAZIONE.

Sia X un sottoinsieme di S_0 .

Ovviamente se l'insieme X è aperto [chiuso], esso è anche aperto [chiuso] nel sottospazio $(S_0 , \mathcal{T}_0)$, perché $X = S_0 \cap X$ e quindi X si esprime come intersezione di S_0 con un aperto [con un chiuso].

Dimostriamo che viceversa, ma <u>solo quando il sostegno S_0 del sottospazio è un insieme aperto [chiuso]</u>, dall'ipotesi che X sia aperto [chiuso] nel sottospazio $(S_0 , \mathcal{T}_0)$ consegue che X è anche aperto [chiuso] .

DIM. Supponiamo S_0 aperto. Se l'insieme X è aperto su S_0, esiste un aperto A tale che $X = S_0 \cap A$ e quindi X, essendo intersezione di due aperti, è un insieme aperto. Se invece S_0 non è aperto, l'implicazione :

(2) $$X \text{ aperto su } S_0 \Rightarrow X \text{ aperto}$$

non sussiste ; infatti , se per assurdo vale la (2) , l'insieme S_0 è necessariamente aperto: ciò si deduce subito utilizzando la (2) con $X = S_0$ (dato che S_0 è aperto su S_0).

Analogamente si prova l'alternativa dell'enunciato, riguardante il caso S_0 chiuso e X chiuso.

---

* * *

Mostriamo che, se $\mathcal{T}$ è la topologia discreta per S (pag.4), nello spazio topologico $(S, \mathcal{T})$ è vera la proposizione seguente:

(3) *Per ogni coppia (x,y) di punti distinti di S esistono un intorno di x e un intorno di y disgiunti.*

DIM. Basta osservare che qualunque sia $x_0 \in S$ l'insieme $\{x_0\}$ è un intorno del punto x_0 , perché con la topologia discreta $\{x_0\}$ è un aperto al quale x_0 appartiene. Pertanto, se x ed y sono punti distinti di S, gli insiemi $\{x\}$ ed $\{y\}$ sono disgiunti, e sono rispettivamente intorni di x ed y. Dunque la (3) è vera.

Invece se $\mathcal{T}$ è la topologia banale , nello spazio topologico $(S,\mathcal{T})$ la proposizione (3) è falsa, perché ciascun punto x_0 di S ha un unico intorno, che è l'insieme S.

Orbene, qualunque sia la topologia $\mathcal{T}$, se accade che nello spazio topologico $(S, \mathcal{T})$ è vera la (3) si dice che $(S,\mathcal{T})$ è uno **spazio di Hausdorff**, oppure uno **spazio separato**. La proposizione (3) viene detta **l'assioma di Hausdorff**.

Ricordiamo che in Analisi Matematica la (3) interviene nella dimostrazione dell'unicità del limite di una successione di numeri reali, o di punti di $\mathbf{R}^n$, nonché nella dimostrazione dell'unicità del limite di una funzione di una o più variabili. Quando introdurremo la nozione di limite per una successione di punti di uno spazio topologico, e per una funzione definita in uno spazio topologico e a valori in uno spazio topologico, vedremo che per dimostrare l'unicità occorre l'ipotesi che lo spazio sia uno spazio di Hausdorff.

E' importante osservare che:

(2.1). *In uno spazio di Hausdorff gli insiemi costituiti da un solo punto sono chiusi; di conseguenza ogni insieme finito è chiuso.*

DIM. Supponiamo che $(S,\mathcal{T})$ sia uno spazio di Hausdorff. Fissato $x \in S$ e detto $\mathcal{I}_c(x)$ l'insieme costituito dagli intorni chiusi di x [non vuoto, dato che S è un intorno di x] , proveremo che risulta:

$$(4) \qquad \bigcap_{I \in \mathcal{I}_c(x)} I = \{x\} \ .$$

Ne seguirà che $\{x\}$ è chiuso, perché la (4) comporta che $\{x\}$ è intersezione di insiemi chiusi.

Siccome x appartiene a qualunque intorno di x, il secondo membro della (4) è incluso nell'insieme a primo membro, e perciò basta dimostrare che quest'ultimo è incluso in $\{x\}$; cioè che un punto y appartenente a tutti gli intorni chiusi di x necessariamente coincide con x. Procedendo per assurdo, ammettiamo che sia $y \neq x$.

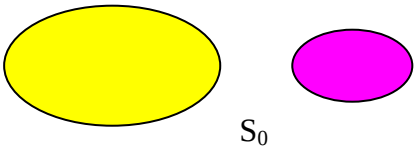
Per l'assioma di Hausdorff esistono allora un intorno I_x di x e un intorno I_y di y disgiunti. Sia A un aperto , al quale appartenga y, contenuto in I_y (esistente per la definizione di intorno). Ciascun punto di I_x non appartiene ad A (in quanto $A \subseteq I_y$ e I_x , I_y sono disgiunti), sicchè $I_x \subseteq -A$ e perciò $-A$, contenendo un intorno di x, è a sua volta un intorno di x, anzi è un intorno chiuso di x (dato che A è aperto). Pertanto $y \in -A$ (in quanto y appartiene a tutti gli intorni chiusi di x) , in contrasto con l'ipotesi fatta che y appartenga ad A.

---

3. Sottoinsiemi connessi di uno spazio topologico; sottoinsiemi compatti e sottoinsiemi precompatti.

Siano $(S, \mathcal{G})$ uno spazio topologico ed S_0 un sottoinsieme non vuoto di S.

Consideriamo una parte X propria e non vuota di S_0: per fissare le idee, si pensi alla parte di colore più chiaro dell'insieme S_0 disegnato qui accanto.



S_0

Diremo che X è **una parte staccata** di S_0 , quando esiste un aperto A contenente X e disgiunto da S_0-X .

Nel caso della figura, S_0-X è la parte di S_0 di colore più scuro, ed è intuitivo che gli insiemi X, S_0-X sono entrambi parti staccate di S_0 : pertanto, con riferimento all'insieme S_0 della figura, esiste una partizione di S_0 , appunto $\{X, S_0\text{-}X\}$, costituita da due parti staccate. Perciò tale insieme non è connesso, conformemente alla definizione che segue.

Diremo che S_0 è un insieme **connesso** quando è vera la proposizione seguente:

Non esiste alcuna partizione di S_0 costituita da due parti staccate .

Di conseguenza, dire che un insieme S_0 *non è connesso* significa che esiste una parte staccata di S_0 , diciamola X, tale che S_0-X sia a sua volta una parte staccata.

E' opportuno ribadire che una parte staccata di un insieme è parte propria e non vuota, ed è importante evidenziare che non basta l'esistenza di una sola parte staccata per dire che un insieme non è connesso : ad esempio si può dimostrare che è connesso l'insieme S_0 unione di due cerchi tangenti esternamente, uno aperto e l'altro chiuso, pur essendo il cerchio aperto una parte staccata di S_0 (la parte rimanente, il cerchio chiuso, non è staccata).

Dimostriamo che:

(3.1) *La proposizione "S_0 è connesso" equivale alla seguente: non esiste alcuna partizione di S_0 costituita da due insiemi non vuoti e aperti su S_0* (*equivalentemente, da due insiemi non vuoti e chiusi su S_0*).

DIM. Conviene dimostrare l'equivalenza tra le negazioni delle due proposizioni.

Supposto S_0 non connesso, sia X una parte staccata di S_0 . Se A è un aperto contenente X e disgiunto da S_0-X, risulta $X = S_0 \cap A$ e quindi X è un insieme aperto su S_0 [pag.9], non vuoto in quanto X è parte staccata di S_0 ; essendo anche l'insieme S_0-X una parte staccata di S_0 , si conclude che la partizione $\{X, S_0\text{-}X\}$ di S_0 è costituita da due insiemi non vuoti e aperti su S_0 . Si noti che i predetti due insiemi, aperti nel sottospazio di S di sostegno S_0 , sono anche chiusi nel sottospazio in quanto ciascuno di essi , nel sottospazio, è il complementare di un aperto.

Viceversa, supponiamo che esista una partizione $\{ X_1 , X_2\}$ di S_0 costituita da due insiemi non vuoti e aperti su S_0. Essendo X_1 aperto su S_0 , esiste un aperto A tale che $X_1 = S_0 \cap A$, sicchè A contiene X_1 ; d'altra parte, dalle relazioni

$$A \cap X_2 \subseteq A \cap S_0 = X_1$$

si deduce che $A \cap X_2 = \varnothing$: infatti se esistesse un punto dell'insieme a primo membro della precedente relazione, esso sarebbe un punto di X_2 e di X_1, il che è assurdo in quanto $X_1 \cap X_2 = \varnothing$.

Dunque l'insieme A contiene X_1 ed è disgiunto da X_2 ; siccome X_1 è non vuoto, ed è diverso da S_0 (altrimenti sarebbe $X_2 = \varnothing$), ne segue che X_1 è una parte staccata di S_0 . Analogamente si conclude che X_2 è una parte staccata di S_0 , e ciò prova che S_0 non è connesso.

Se l'insieme S_0 è aperto [chiuso], dire che è connesso equivale a dire che non si può esprimere come unione di due aperti [chiusi] disgiunti e non vuoti : ciò si deduce dalla (3.1), in base all'Osservazione di pag.10.

E' importante approfondire l'eventualità che sia connesso il sostegno S dello spazio topologico $(S,\mathcal{G})$, nel qual caso si dice brevemente che è connesso *lo spazio S.* Il teorema che segue fornisce diverse caratterizzazioni di tale proprietà : è più semplice, come noi faremo, formularlo e dimostrarlo con riferimento alla negazione della proprietà stessa.

(3.2). *Le proposizioni che seguono sono equivalenti:*

a) *lo spazio S non è connesso.*

b) *esiste una partizione di S costituita da due aperti non vuoti.*

c) *esiste una partizione di S costituita da due chiusi non vuoti.*

d) *esistono sottoinsiemi di S, propri e non vuoti, simultaneamente aperti e chiusi.*

e) *esistono sottoinsiemi di S, propri e non vuoti, la cui frontiera è l'insieme vuoto.*

DIM. Siccome *aperto su* $S \Leftrightarrow$ *aperto* , dalla (3.1) con $S_0=S$ consegue che a) $\Leftrightarrow$ b) $\Leftrightarrow$ c). Supponiamo ora che sia vera la b) : se $\{ X_1 , X_2 \}$ è una partizione di S costituita da due insiemi aperti e non vuoti, ciascuno dei due è anche chiuso (in quanto complementare di un aperto), ed è anche diverso da S (altrimenti l'altro sarebbe vuoto) : pertanto b) $\Rightarrow$ d). Per provare che d) $\Rightarrow$ b), supponiamo che esista un sottoinsieme X di S, proprio e non vuoto, simultaneamente aperto e chiuso : allora anche S-X è simultaneamente aperto e chiuso , ed è a sua volta proprio e non vuoto (se fosse S-$X = S$ [$= \varnothing$], sarebbe $X = \varnothing$ [$= S$]) : ne segue che $\{X, S\text{-}X\}$ è una partizione di S costituita da due aperti non vuoti, e pertanto è vera la b).

Resta da stabilire che la e) è equivalente ad una delle proposizioni precedenti : orbene, dalla (6) di pag.8 segue subito che d) $\Leftrightarrow$ e).

* * *

L'importante nozione di *insieme compatto* , che in $\mathbf{R}^n$ equivale per definizione a quella di *insieme chiuso e limitato*, si estende ai sottoinsiemi di uno spazio topologico $(S,\mathcal{G})$ ma non negli stessi termini : ciò , com'è ovvio, perché in uno spazio topologico qualunque non ha senso parlare di insiemi limitati ; peraltro questo non è il solo motivo, dato che in certi spazi topologici di tipo particolare si introduce, come vedremo, la nozione di *insieme limitato*, ma in tali spazi la proprietà *chiuso e limitato* in generale non equivale a *compatto.*

Prima di introdurre la predetta nozione, è opportuno ricordare quella di *ricoprimento* di un sottoinsieme X di un insieme S.

Una famiglia di sottoinsiemi di S (non necessariamente a due a due disgiunti), diciamola $\mathcal{R} = (R_i)_{i\in I}$, si chiama un *ricoprimento* di X quando l'unione degli insiemi della famiglia contiene X ; in simboli, quando risulta:

$$\bigcup_{i\in I} R_i \supseteq X \ .$$

Aggiungiamo qui che se l'unione di "alcuni " degli insiemi della famiglia $\mathcal{R}$ contiene ancora l'insieme X, essi costituiscono una "sottofamiglia" di $\mathcal{R}$, diciamola $\mathcal{R}_0$, che è a sua volta un ricoprimento di X : quest'ultimo si può chiamare un *sottoricoprimento* di X, rispetto ad $\mathcal{R}$.

Formuliamo il significato di tale locuzione in modo preciso.

Sia I_0 un sottoinsieme proprio e non vuoto di I, cioè del cosiddetto *insieme degli indici* della famiglia $\mathcal{R}$. Se accade che la famiglia $\mathcal{R}_0=(R_i)_{i\in I_0}$ è a sua volta un ricoprimento di X , essa si chiama un **sottoricoprimento** di X (**rispetto ad** $\mathcal{R}$) ; brevemente, si dice che $\mathcal{R}$ **contiene il ricoprimento** $\mathcal{R}_0$. Quando I_0 è un insieme finito,

sicchè gli insiemi della famiglia $\mathcal{R}_0$ sono in numero finito, si dice che $\mathcal{R}_0$ è un **ricoprimento finito** di X.

Non è detto che un dato ricoprimento $\mathcal{R}$ dell'insieme X contenga dei ricoprimenti di X ; se ne contiene, quando l'insieme degli indici di $\mathcal{R}$ è infinito può accadere che nessuno dei sottoricoprimenti sia finito.

Fin qui ci siamo riferiti ad un insieme S, dato che non occorreva una topologia. Supponiamo ora che S sia il sostegno di uno spazio topologico.

Se gli insiemi del ricoprimento $\mathcal{R}$ di X sono aperti [aperti su X] si dice che la famiglia $\mathcal{R}$ è un **ricoprimento aperto** [**aperto su** X] dell'insieme X.

Ciò premesso, un sottoinsieme X di uno spazio topologico $(S,\mathcal{G})$ si dice **compatto** quando è vera la proposizione seguente:

Ogni ricoprimento aperto di X contiene un ricoprimento finito di X ,

nella quale la parola "*aperto*" si può sostituire con "*aperto su X*". Infatti si ha che:

(3.3). *L'insieme X è compatto se e solo se ogni ricoprimento di X , aperto su X, contiene un ricoprimento finito di X.*

La dimostrazione non presenta alcuna difficoltà, ma non è il caso di riportarla.

Abbiamo già detto che la nozione di insieme compatto estende agli spazi topologici quella che si introduce in $\mathbf{R}^n$ [munito della topologia usuale] : ciò risulta dal seguente teorema che ci limitiamo ad enunciare.

TEOREMA DI HEINE-PINCHERLE-BOREL.

Nello spazio topologico di sostegno $\mathbf{R}^n$ *(con la topologia usuale), un insieme è compatto se e solo se è chiuso e limitato.*

Dimostriamo che:

(3.4). *In uno spazio topologico di Hausdorff ogni insieme compatto è chiuso.*

DIM. Sia X un insieme compatto. Essendo $X \subseteq \overline{X}$ [(1) pag.6], in virtù della (1.3) di pag.6 per provare che X è chiuso basta dimostrare che sussiste l'inclusione opposta $\overline{X} \subseteq X$, cioè (passando ai complementari) che risulta:

(1) $$-X \subseteq -\overline{X} \quad .$$

Sia dunque y un punto qualunque di $-X$.

Per l'assioma di Hausdorff a ciascun punto $x \in X$ (il quale è diverso da y) si possono associare un intorno I_x di x ed un intorno J_x di y che siano disgiunti. Dunque risulta:

(2) $$I_x \cap J_x = \emptyset \qquad \forall\, x \in X \quad .$$

Per ogni $x \in X$ possiamo supporre che l'intorno I_x sia aperto , sostituendolo in caso contrario con un aperto in esso contenuto al quale x appartenga.

Siccome ovviamente $X \subseteq \bigcup_{x\in X} I_x$, la famiglia $\mathcal{R} = (I_x)_{x\in X}$ è un ricoprimento aperto di X: essendo X compatto, $\mathcal{R}$ contiene un ricoprimento finito di X , diciamolo $\{I_{x_1}, I_{x_2},..., I_{x_k}\}$. Si ha dunque:

(3) $$X \subseteq I_{x_1} \cup I_{x_2} \cup ... \cup I_{x_k} .$$

D'altra parte, poiché l'intersezione di due intorni di un punto, e quindi anche l'intersezione di un numero finito di intorni, è a sua volta un intorno del punto, l'insieme

(4) $$J = J_{x_1} \cap J_{x_2} \cap ... \cap J_{x_k}$$

è un intorno di y che, come andiamo a dimostrare, non ha punti in comune con X.

Infatti, procedendo per assurdo, ammettiamo che esista un punto x_0 di J appartenente ad X, quindi per la (3) appartenente ad almeno uno degli insiemi a secondo membro della (3), diciamolo I_{x_r} con $r \in \{1, 2, \ldots, k\}$. Siccome per la (4) il punto y , in quanto punto di J , deve appartenere a ciascuno degli insiemi a secondo membro della (4), si ha $x_0 \in I_{x_r} \cap J_{x_r}$ in contrasto con la (2).

Poiché esiste un intorno di y, precisamente J, al quale non appartiene alcun punto di X, si conclude che y non è aderente ad X, cioè che y appartiene all'insieme $-\overline{X}$.

Dopo ciò la (1) è dimostrata, e con essa il teorema.

Per il teorema ora acquisito, in uno spazio di Hausdorff un insieme X non può essere compatto se non è chiuso; peraltro può accadere che sia compatto l'insieme $\overline{X}$ [dato che è chiuso, per la (1.2) di pag.6]. Orbene, con riferimento ad un sottoinsieme *X di un qualunque spazio topologico*, se la chiusura di un insieme X è compatta, si dice che X è **precompatto**, o che X è **relativamente compatto**.

Tenendo presente il teorema di Heine-Pincherle-Borel (pag.14), è facile constatare che:

(3.5). *I sottoinsiemi precompatti di* $\mathbf{R}^n$ *sono quelli limitati* .

DIM. In $\mathbf{R}^n$ si ha:

X *insieme limitato* $\Leftrightarrow$ $\overline{X}$ *insieme limitato* $\Leftrightarrow$ $\overline{X}$ *insieme chiuso e limitato* $\Leftrightarrow$ $\overline{X}$ *insieme compatto.*

Nello spazio topologico $\mathbf{R}^n$ [cfr. inizio pag.9] le nozioni di *insieme precompatto* e di *insieme compatto* si possono esprimere mediante una proprietà delle successioni di punti dell'insieme. Se $X \subseteq \mathbf{R}^n$ è un insieme precompatto, ossia limitato [per la (3.5)], ogni successione di punti di X è limitata e quindi da essa, com'è noto, si può estrarre una successione convergente; viceversa è facile dimostrare [con un ragionamento per assurdo], che se da ogni successione di punti di X si può estrarre una successione convergente, allora l'insieme X è limitato, ossia precompatto. Dunque nello spazio topologico $\mathbf{R}^n$ dire che "*l' insieme X è precompatto*" equivale a dire che "*da ogni successione di punti di X si può estrarre una successione convergente*".

Da ciò si deduce che, nello spazio topologico $\mathbf{R}^n$, la proposizione "*l'insieme X è compatto*" equivale alla seguente: " *da ogni successione di punti di X si può estrarre una successione convergente ad un punto di X* " . Omettiamo la dimostrazione, perché sulle caratterizzazioni appena esposte torneremo nel seguito : esse sussistono non solo in $\mathbf{R}^n$, ma in tutti gli spazi topologici chiamati spazi *metrici* (saranno introdotti nel cap.IV).

Concludiamo questo numero con la seguente definizione.

Uno spazio topologico $(S, \mathcal{T})$ si dice **localmente compatto** quando per ogni $x \in S$ esiste un intorno di x che è un insieme compatto.

4. Assiomi di Hausdorff. Spazio topologico prodotto.

Gli *assiomi di Hausdorff*, che ci proponiamo di formulare in questo numero, consentono di " costruire " uno spazio topologico di sostegno S scegliendo , invece che una classe di sottoinsiemi di S destinata a costituire la classe degli aperti, una classe di insiemi $\mathcal{J}(x)$, per ogni $x \in S$, destinata ad essere la classe degli intorni di x.

Sia S un insieme, e per ogni elemento x di S sia $\mathcal{J}(x)$ una classe non vuota di sottoinsiemi di S diversi da Ø.

Supponiamo che per la famiglia $(\mathcal{J}(x))_{x\in S}$ si abbia quanto segue:

Qualunque sia l'elemento x di S sono verificate le seguenti condizioni:

h$_1$) *x appartiene ad ogni insieme della classe* $\mathcal{J}(x)$

h$_2$) $I \in \mathcal{J}(x)$ *e* $J \supseteq I \Rightarrow J \in \mathcal{J}(x)$

h$_3$) $I_1 \in \mathcal{J}(x)$ e $I_2 \in \mathcal{J}(x) \Rightarrow I_1 \cap I_2 \in \mathcal{J}(x)$

h$_4$) *Ogni insieme I della classe* $\mathcal{J}(x)$ *contiene un sottoinsieme A di S, al quale x appartenga, che verifichi l'equivalenza*:

(1) $$a \in A \Leftrightarrow A \in \mathcal{J}(a) \ .$$

Si dice allora che la famiglia $(\mathcal{J}(x))_{x\in S}$ verifica **gli assiomi di Hausdorff** : così si chiamano, infatti, le condizioni esposte con l'intesa che siano verificate per ogni $x \in S$.

Sussiste il teorema:

(4.1). *Si supponga che la famiglia* $(\mathcal{J}(x))_{x\in S}$ *verifichi gli assiomi di Hausdorff , e si denoti con* $\mathcal{T}$ *l'insieme costituito da Ø e dalle parti A di S che verificano l'equivalenza* (1).

Allora $\mathcal{T}$ *è una topologia per S, e nello spazio topologico* $(S,\mathcal{T})$ *gli insiemi della classe* $\mathcal{J}(x)$*, per ogni* $x \in S$*, sono gli intorni del punto x.*

La topologia $\mathcal{T}$ *è l'unica soddisfacente a tale condizione, e sarà detta la topologia per S* **determinata dalla famiglia** $(\mathcal{J}(x))_{x\in S}$ (**famiglia degli intorni**).

DIM. In primo luogo occorre dimostrare che $\mathcal{T}$ verifica le condizioni 1,2,3 di pag.4.

1. L'insieme Ø appartiene a $\mathcal{T}$ per definizione; l'insieme S appartiene a $\mathcal{T}$ in quanto per $A=S$ le due proposizioni a sinistra e a destra della (1) sono entrambe vere : la prima banalmente, la seconda perché, scelto ad arbitrio un elemento I di $\mathcal{J}(a)$ [esistente per l'ipotesi $\mathcal{J}(x) \neq \emptyset$] , dalla h$_2$ con $x=a$ *e* $J=S$ si deduce $S \in \mathcal{J}(a)$.

2. Siano A_1 ed A_2 elementi non vuoti di $\mathcal{T}$, sicchè $a \in A_1 \Leftrightarrow A_1 \in \mathcal{J}(a)$ e $a \in A_2 \Leftrightarrow A_2 \in \mathcal{J}(a)$. Allora, tenuto conto della h$_3$, sussistono le equivalenze: $a \in A_1 \cap A_2 \Leftrightarrow [\, A_1 \in \mathcal{J}(a)$ e $A_2 \in \mathcal{J}(a)\,] \Leftrightarrow A_1 \cap A_2 \in \mathcal{J}(a)$, e quindi è vera la (1) con $A = A_1 \cap A_2$, ossia $A_1 \cap A_2$ è un elemento di $\mathcal{T}$. Se uno degli insiemi A_1 ed A_2 è vuoto, è banale che $A_1 \cap A_2 \in \mathcal{T}$.

3. Sia $(A_i)_{i\in I}$ una famiglia di elementi di $\mathcal{T}$, sicchè se $A_r \neq \emptyset$ è vera l'equivalenza

(*) $$a \in A_r \Leftrightarrow A_r \in \mathcal{J}(a) \qquad \forall\, r \in I \ .$$

Si tratta di dimostrare che:

$$a \in \bigcup_{i\in I} A_i \Leftrightarrow \bigcup_{i\in I} A_i \in \mathcal{J}(a) \ .$$

Si ha infatti, tenendo presente la (*) :

$$a \in \bigcup_{i\in I} A_i \Leftrightarrow \exists r \in I : a \in A_r \Leftrightarrow \exists r \in I : A_r \in \mathcal{J}(a) \Leftrightarrow \bigcup_{i\in I} A_i \in \mathcal{J}(a) \ .$$

Abbiamo acquisito che $\mathcal{T}$ è una topologia per S . Mostriamo ora che per ogni $x \in S$ nello spazio topologico $(S\ ,\ \mathcal{T})$ gli insiemi della famiglia $(\mathcal{J}(x))_{x\in S}$ sono gli intorni del punto x, cioè che per ogni $x \in S$ si ha:

(2) $$I \text{ intorno di } x \text{ in } (S,\mathcal{T}) \Leftrightarrow I \in \mathcal{J}(x) \ .$$

Se I è un intorno di x, I contiene un aperto A al quale x appartiene, ossia tale che $A \in \mathcal{I}(x)$ [siccome $A \in \mathcal{T}$, vale la (1) con $a=x$] : per la condizione h_2 ne segue $I \in \mathcal{I}(x)$ [dato che $A \in \mathcal{I}(x)$ e $I \supseteq A$]. Viceversa, supposto che $I \in \mathcal{I}(x)$, poniamo:

$$B = \{a \in S : I \in \mathcal{I}(a)\}.$$

Ovviamente $x \in B$, e d'altra parte $B \subseteq I$ [se $b \in B$, cioè se $I \in \mathcal{I}(b)$, per la h_1 risulta $b \in I$] . Per concludere che I è un intorno di x basta dunque mostrare che B è un aperto; ciò vuol dire che B è un elemento di $\mathcal{T}$, ossia che risulta:

$$b \in B \Leftrightarrow B \in \mathcal{I}(b) \ .$$

Orbene, se $b \in B$, cioè se $I \in \mathcal{I}(b)$, per la h_4 esiste $A \subseteq B$ soddisfacente alla (1) tale che $b \in A$: allora $A \in \mathcal{I}(b)$ [per la (1) con $a=b$], e quindi per la h_2 si ha $B \in \mathcal{I}(b)$. Viceversa se $B \in \mathcal{I}(b)$, per la h_1 si ha $b \in B$.

Resta da dimostrare che $\mathcal{T}$ è l'unica topologia per S tale che per ogni $x \in S$ sussista la (2). Se $\mathcal{T}_1$ è una siffatta tolologia, cioè una topologia tale che per ogni $x \in S$ si ha

(3) $$I \text{ intorno di } x \text{ in } (S,\mathcal{T}_1) \Leftrightarrow I \in \mathcal{I}(x) \quad ,$$

dalle (2),(3) consegue, per ogni $x \in S$:

$$I \text{ intorno di } x \text{ in } (S,\mathcal{T}) \Leftrightarrow I \text{ intorno di } x \text{ in } (S, \mathcal{T}_1).$$

In particolare , sussiste l'equivalenza:

$$I \text{ intorno di } x \text{ in } (S,\mathcal{T}) \ \forall\, x \in I \Leftrightarrow I \text{ intorno di } x \text{ in } (S, \mathcal{T}_1) \ \forall\, x \in I \ ,$$

ossia, per la (1.1) di pag.5:

$$I \text{ aperto in } (S,\mathcal{T}) \Leftrightarrow I \text{ aperto in } (S, \mathcal{T}_1).$$

Tale equivalenza significa che $\mathcal{T} = \mathcal{T}_1$, e l'asserto è provato .

Ad esempio per ogni $x \in \mathbf{R}^n$ sia $\mathcal{I}(x)$ la classe dei sottoinsiemi di $\mathbf{R}^n$ ciascuno dei quali contiene un intervallo aperto [un cerchio aperto] di centro x : è facile constatare che in entrambe le alternative la famiglia $(\mathcal{I}(x))_{x \in S}$ verifica gli assiomi di Hausdorff, e quindi per la (4.1) determina una topologia per $\mathbf{R}^n$: ovviamente questa è la stessa nelle due alternative, e coincide con la topologia usuale per $\mathbf{R}^n$ [cfr. inizio pag.9]. Un ulteriore esempio sull'uso della (4.1) sarà esposto a pag.30 [es.6].

La (4.1) si utilizza altresì per introdurre la nozione che andiamo ad esporre, quella di *spazio topologico prodotto.*

Consideriamo una ennupla di spazi topologici :

(3) $$\big((S_1,\mathcal{T}_1), \ldots, (S_n, \mathcal{T}_n)\big),$$

e poniamo $S = S_1 \times S_2 \times \ldots \times S_n$. Vogliamo costruire uno spazio topologico di sostegno S in modo che si abbia quanto segue: per ogni $x = (x_1,x_2,\ldots,x_n) \in S$, un intorno I di x sia un qualunque sottoinsieme di S che contenga un insieme del tipo $I_1 \times I_2 \times \ldots \times I_n$, dove I_1, I_2, …., I_n sono intorni rispettivamente di x_1, x_2,…,x_n nei rispettivi spazi topologici $(S_1,\mathcal{T}_1)$, $(S_2,\mathcal{T}_2)$, ….., $(S_n,\mathcal{T}_n)$.

Orbene, denotata con $\mathcal{I}(x)$ la classe dei sottoinsiemi I di S del tipo appena precisato, la famiglia $(\mathcal{I}(x))_{x \in S}$ verifica gli assiomi di Hausdorff [omettiamo la dimostrazione, che costituisce un facile esercizio per il lettore] : la topologia $\mathcal{T}$ determinata dalla predetta famiglia si chiama **prodotto delle topologie** $\mathcal{T}_1$, $\mathcal{T}_2$, …, $\mathcal{T}_n$, e lo spazio topologico $(S,\mathcal{T})$ si chiama **prodotto degli spazi topologici** della ennupla (3).

5. Basi di topologia.

Sia S un insieme, e sia $\mathcal{B}$ una classe di sottoinsiemi di S .

Consideriamo l'insieme $\mathcal{T}$ costituito dalle parti di S ciascuna delle quali si può esprimere come unione di elementi di $\mathcal{B}$: se $\mathcal{T}$ è una topologia per S, si dice che la classe $\mathcal{B}$ è **una base di topologia (**per S), e quindi che $\mathcal{T}$ è **la topologia per** S **di base** $\mathcal{B}$**.** Diremo pure che $\mathcal{T}$ **è dotata della base** $\mathcal{B}$, o che $\mathcal{B}$ **è una base di** $\mathcal{T}$.

Ad esempio la classe $\mathcal{B}$ degli intervalli aperti di $\mathbf{R}^n$, e quella dei cerchi aperti di $\mathbf{R}^n$, sono basi della topologia usuale per $\mathbf{R}^n$ [cfr. inizio pag.9] .

Uno spazio topologico di sostegno S si può "costruire" anche scegliendo, invece che una classe di sottoinsiemi di S destinata a costituire la classe degli aperti, una classe $\mathcal{B}$ che sia una base di topologia: dopo ciò, all'insieme S si associa la topologia di base $\mathcal{B}$.

Il teorema che segue fornisce delle condizioni affinchè una data classe $\mathcal{B}$ sia una base di topologia.

(5.1). *Sia* $\mathcal{B}$ *una classe di sottoinsiemi di* S *soddisfacente alle seguenti condizioni*:

a). $\emptyset \in \mathcal{B}$, *e* S *è esprimibile come unione di elementi di* $\mathcal{B}$.

b). *Qualunque siano* $B_1 , B_2 \in \mathcal{B}$ *tali che* $B_1 \cap B_2 \neq \emptyset$, *per ogni* $x \in B_1 \cap B_2$ *esiste un insieme* $B_x \subseteq B_1 \cap B_2$ *della classe* $\mathcal{B}$ *al quale* x *appartiene.*

Allora la classe $\mathcal{B}$ *è una base di topologia*.

DIM. Sia $\mathcal{T}$ l'insieme costituito dalle parti di S ciascuna delle quali si può esprimere come unione di elementi di $\mathcal{B}$. Per provare che $\mathcal{T}$ è una topologia per S , si tratta di stabilire che sono soddisfatte le condizioni 2,3 di pag,4, dato che la 1 è conseguenza banale della condizione a) dell'enunciato.

Che sia soddisfatta la 3 è evidente: qualunque sia la famiglia $(A_i)_{i\in I}$ di elementi di $\mathcal{T}$, l'insieme $\bigcup_{i\in I} A_i$ è unione di unioni di elementi di $\mathcal{B}$, sicchè è a sua volta unione di elementi di $\mathcal{B}$ e pertanto appartiene a $\mathcal{T}$.

Per dimostrare che è soddisfatta la 2 di pag.4 osserviamo intanto che dalla condizione b), con le notazioni ivi adoperate, consegue l'uguaglianza:

$$B_1 \cap B_2 = \bigcup_{x\in B_1\cap B_2} B_x$$

e quindi l'intersezione di due insiemi della classe $\mathcal{B}$, anche se non è vuota, si può esprimere come unione di insiemi della stessa classe. Ciò premesso, siano A_1, A_2 due elementi di $\mathcal{T}$:

$$A_1 = \bigcup_{i\in I} B_{1,i} \quad , \quad A_2 = \bigcup_{j\in J} B_{2,j}$$

con $B_{1,i}$, $B_{2,j}$ elementi di $\mathcal{B}$. Dalla relazione:

(1) $$A_1 \cap A_2 = \bigcup_{i\in I\,,\, j\in J} B_{1,i} \cap B_{2,j} \quad ,$$

tenuto conto di quanto dianzi osservato, emerge che $A_1 \cap A_2$ si può esprimere come unione di unioni di elementi di $\mathcal{B}$, e quindi appartiene a $\mathcal{T}$.

Se $\mathcal{B}$ è una base di topologia finita o numerabile, lo spazio topologico di base $\mathcal{B}$ gode di una importante proprietà che vedremo nel numero seguente.

6. Spazi separabili.

Consideriamo uno spazio topologico S .

Un sottoinsieme Δ di S si dice **denso in** S [alcuni adoperano la locuzione **ovunque denso in** S] , quando risulta $\overline{\Delta} = S$, ossia quando ogni punto di S è aderente a Δ : in altri termini quando, qualunque sia $x \in S$, accade che a ciascun intorno di x appartiene almeno un punto di Δ .

Quando esiste un insieme finito o numerabile denso in S, lo spazio topologico si dice **separabile**.

Ad esempio l'insieme **Q** dei numeri razionali è denso in **R** , inteso come spazio topologico con la topologia usuale [cfr. inizio pag.9]. Siccome **Q** è un insieme numerabile, lo spazio **R** è separabile.

Non si confonda la nozione dianzi introdotta con quella di *spazio separato* o *spazio di Hausdorff* (pag.11).

Sussiste il teorema:

(6.1). *Se la topologia di uno spazio topologico S é dotata di una base* $\mathcal{B}$ *finita o numerabile, allora lo spazio S è separabile.*

DIM. In ciascun elemento non vuoto B della base $\mathcal{B}$ scegliamo un punto ; denotiamo quindi con Δ l'insieme dei punti scelti: essendo $\mathcal{B}$ finito o numerabile, tale è anche Δ .

Detto x un qualunque punto di S, ciascun intorno I di x contiene un aperto A al quale x appartiene, e d'altra parte A, essendo unione di elementi di $\mathcal{B}$, contiene almeno un $B \in \mathcal{B}$ diverso da Ø : poiché all'insieme B appartiene il punto in esso prescelto , elemento di Δ , tale punto appartiene anche ad I.

Dunque a ciascun intorno I di x appartiene almeno un punto di Δ e quindi , essendo x arbitrario, l'insieme Δ è denso in S. Siccome Δ è finito o numerabile l'asserto è provato.

---

7. Limite di una successione. Limite e continuità di una funzione in un punto.

Sia $(x_n)_{n \in N}$ una successione di punti di uno spazio topologico S. Un punto x_0 di S si chiama **limite** della successione quando è vera la proposizione seguente:

(1) $$\forall\, I \text{ intorno di } x_0,\ \exists\, \nu \in \mathbf{N} : \quad n > \nu \Rightarrow x_n \in I \ .$$

Se esiste almeno un punto x_0 per il quale è vera la (1) si dice che la successione è **convergente** , e che essa **converge ad** x_0.

Quando la topologia dello spazio è quella banale, la (1) è vera per qualunque successione e per qualunque punto x_0 dello spazio [dato che x_0 ha un unico intorno, l'insieme S] : perciò in tale spazio tutte le successioni sono convergenti, e qualunque successione converge ad un qualunque punto.

E' facile costruire esempi di spazi topologici , con la topologia diversa da quella banale, nei quali esistono successioni convergenti che non hanno un unico limite [si veda l'es.3 di pag.28]. In proposito sussiste il teorema:

(7.1) *Se S è uno spazio di Hausdorff, ogni successione convergente di punti di S ha un unico limite.*

DIM. Procedendo per assurdo, ammettiamo che la successione $(x_n)_{n\in N}$ converga simultaneamente a due punti distinti x_0, y_0 di *S*. Per l'assioma di Hausdorff esistono un intorno *I* di x_0 e un intorno *J* di y_0 tali che $I \cap J = \emptyset$; siano $\nu, \nu' \in \mathbf{N}$ tali che si abbia:

$$x_n \in I \quad \forall\ n > \nu \quad , \quad x_n \in J \quad \forall\ n > \nu' \ .$$

Scelto $n \in \mathbf{N}$ maggiore sia di ν che di ν' , ne segue che $x_n \in I \cap J$ e ciò è assurdo in quanto $I \cap J = \emptyset$.

Rileviamo che l'ipotesi "*spazio di Hausdorff* " non è necessaria per l'unicità del limite : ossia esistono spazi che non sono di Hausdorff nei quali ogni successione convergente ha un unico limite [cfr. pag.28.1, in particolare pag.28.2].

Se *S* è uno spazio di Hausdorff, l'unico limite della successione convergente $(x_n)_{n\in N}$ si denota con uno dei simboli:

$$\lim_{n\to\infty} x_n \ , \ \lim x_n \ ,$$

dei quali è preferibile il secondo quando non vi è possibilità di equivoci.

* * *

Nei discorsi che seguono, *S* ed S' denotano spazi topologici .

Consideriamo una funzione *f* definita in un sottoinsieme *X* di *S* ed a valori in *S*; indichiamo quindi con x_0 un punto di accumulazione per *X* , e con y_0 un punto di S'.

Si dice che y_0 è **limite di** *f* **nel punto** x_0 , oppure che *f*(*x*) **tende ad** y_0 **per** *x* **tendente ad** x_0 , quando è vera la proposizione seguente:

$$\forall\, I' \textit{ intorno di } y_0\,,\ \exists\, I \textit{ intorno di } x_0 \textit{ tale che si abbia}:$$

(*) $$x \in I \cap X - \{x_0\} \ \Rightarrow\ f(x) \in I' \ .$$

Invece di dire "$\exists$ *I intorno di* x_0 " , si può equivalentemente dire "$\exists$ *J intorno di* x_0 *su X* ", e sostituire nella precedente implicazione $I \cap X$ con *J*.

Se esiste almeno un limite di f in x_0 , si dice che *f* è **convergente nel punto** x_0.

OSSERVAZIONE.

L'esistenza di un punto $y_0 \in S'$ limite di f in x_0 , è una proprietà di f che esprime il comportamento della funzione " nelle vicinanze " di x_0 , ma non coinvolge il valore di *f* in x_0 (se x_0 appartiene ad *X*) : infatti l'implicazione (*) impone l'appartenenza di *f*(*x*) all'intorno I' quando *x* , punto di $I \cap X$, è diverso da x_0 . Ovviamente quando x_0 non appartiene ad *X* , nella (*) si può equivalentemente sostituire $I \cap X - \{x_0\}$ con $I \cap X$ (dato che quest'ultimo insieme coincide con il precedente).

In conseguenza : quando $x_0 \in X$, se si considera una funzione ottenuta da *f* modificando solo il valore in x_0 , accade che il punto y_0 è anche limite in x_0 di tale funzione; quando $x_0 \notin X$, se si prolunga la *f* su $X \cup \{x_0\}$ associando ad x_0 un punto qualunque dello spazio S', accade che il punto y_0 è anche limite in x_0 del prolungamento considerato.

Come per le successioni, può accadere che esistano più limiti di f nel punto x_0; con ragionamento analogo a quello esposto per le successioni si dimostra che:

(7.2). *Se S' è uno spazio di Hausdorff, $f(x)$ non può tendere a due limiti distinti per x tendente ad x_0 .*

Quando S' è uno spazio di Hausrdorff, se f è convergente nel punto x_0 l'unico limite si denota col simbolo:

$$\lim_{x\to x_0} f(x) .$$

Per la nozione di limite di f in un punto x_0 non occorre che x_0 appartenga all'insieme di definizione X di f, ma è necessario che x_0 sia di accumulazione per X . Invece per la nozione di *continuità nel punto* x_0 , che ci accingiamo ad introdurre con riferimento alla considerata funzione f , non occorre che x_0 sia di accumulazione per X , ma è necessario che x_0 appartenga ad X.

Sia dunque x_0 un punto di X, e si supponga che S' sia uno spazio di Hausdorff.

Si dice che f è **continua nel punto** x_0 quando è vera la proposizione seguente:

(2) $\quad \forall\ I'$ *intorno di* $f(x_0)$, $\exists\ I$ *intorno di* x_0 *tale che si abbia*:

$$x\in I\cap X \ \Rightarrow f(x)\in I'$$

Anche qui , invece di "$\exists\ I$ *intorno di* x_0 " si può equivalentemente dire "$\exists\ J$ *intorno di* x_0 *su* X ", e sostituire nella precedente implicazione $I\cap X$ con J.

Notiamo che se x_0 è un punto isolato di X la (1) è vera qualunque sia la funzione f definita in X , sicchè *in un eventuale punto isolato dell'insieme di definizione, qualunque funzione è continua.* Invece, come si verifica facilmente, se il punto x_0 dell'insieme X è di accumulazione per X, si ha:

$$f \text{ continua in } x_0 \Leftrightarrow \lim_{x\to x_0} f(x)=f(x_0) .$$

Si dice che la funzione f è **continua in un insieme** $B\subseteq X$, quando essa è continua in ciascun punto $x_0\in B$. Se $B=X$ si dice semplicemente che f **è continua**.

Come nel caso delle funzioni reali di una variabile reale si dimostra che:

(7.3). *Una funzione composta mediante funzioni continue è a sua volta una funzione continua.*

* * *

Prima di riportare ulteriori teoremi sulle funzioni continue, ci sembra utile esporre alcuni richiami di teoria degli insiemi con riferimento ad una funzione f: $A \to B$.

Qualunque sia $Y\subseteq B$, il sottoinsieme di A:

$$\{x\in A : f(x) \in Y\}$$

si denota col simbolo $f^{-1}(Y)$ e si chiama **la controimmagine di** Y **mediante** f . Avvertiamo che in tale simbolo la notazione f^{-1} non sta ad indicare la funzione inversa di f, perché non si suppone che f sia invertibile. Peraltro, qualora f sia invertibile e l'insieme Y sia contenuto nel codominio di f, la notazione f^{-1} rappresenta proprio l'inversa di f , e la controimmagine $f^{-1}(Y)$ coincide con l'immagine di Y mediante f^{-1} .

In proposito ci saranno utili le ovvie osservazioni che seguono.

Per $Y \subseteq B$ sussiste l'equivalenza:

(3) $$x \in f^{-1}(Y) \quad \Leftrightarrow \quad f(x) \in Y \ .$$

Qualunque siano i sottoinsiemi Y_1, Y_2 di B, si ha:

(4) $$f^{-1}(Y_1 - Y_2) = f^{-1}(Y_1) - f^{-1}(Y_2) \ , \qquad Y_1 \subseteq Y_2 \Rightarrow f^{-1}(Y_1) \subseteq f^{-1}(Y_2) \ .$$

Qualunque siano $X_0 \subseteq X$, $Y \subseteq B$ sussistono le inclusioni:

(5) $$f^{-1}\big(f(X_0)\big) \supseteq X_0 \quad , \qquad f\big(f^{-1}(Y)\big) \subseteq Y$$

e vale l'uguaglianza nella prima se $X_0 = X$, nella seconda se $Y \subseteq f(X)$.

In generale nelle (5) l'uguaglianza non sussiste: basta considerare la funzione $f : x \in \mathbf{R} \to \operatorname{sen} x$ e porre

$$X_0 = [0, \pi/2] \ , \ Y = [-2, 2] \ .$$

Si ha $f(X_0) = [0,1]$ e $f^{-1}([0,1]) \supset [0, \pi/2]$, sicché $f^{-1}\big(f(X_0)\big)$ non è uguale ad X_0 ; inoltre $f^{-1}(Y) = \mathbf{R}$ e $f(\mathbf{R}) = [-1,1]$, sicché $f\big(f^{-1}(Y)\big)$ non è uguale ad Y.

* * *

Dimostriamo i teoremi seguenti.

(7.4). *Per ogni funzione* $f : X \subseteq S \to S'$ *le proposizioni che seguono sono equivalenti*:

a) *f è continua.*

b) *per ogni sottoinsieme A aperto su* $f(X)$ *oppure aperto* (*in* S'), *l'insieme* $f^{-1}(A)$ *è aperto su X.*

c) *per ogni sottoinsieme C chiuso su* $f(X)$ *oppure chiuso* (*in* S'), *l'insieme* $f^{-1}(C)$ *è chiuso su X.*

DIM.

a) $\Rightarrow$ b). Fissato un insieme A aperto su $f(X)$, sia x_0 un punto di $f^{-1}(A)$, cioè per la (3) tale che $f(x_0) \in A$. Essendo A aperto su $f(X)$, A è un intorno di $f(x_0)$ su X : per la continuità di f in x_0 esiste quindi un intorno J di x_0 su X tale che si abbia:

$$x \in J \Rightarrow f(x) \in A \ , \quad \text{ossia per la (3)} \quad x \in J \Rightarrow x \in f^{-1}(A) \ ;$$

pertanto $f^{-1}(A) \supseteq J$ e quindi $f^{-1}(A)$ è un intorno di x_0 su X. Dunque l'insieme $f^{-1}(A)$ è intorno su X di ogni suo punto, e quindi è aperto su X [per la (1.1) di pag.5 applicata al sottospazio di S di sostegno X] .

b) $\Rightarrow$ a). Sia x_0 un punto di X, e sia I' un intorno di $f(x_0)$ su $f(X)$, sicchè I' contiene un insieme A aperto su $f(X)$ a cui $f(x_0)$ appartiene: dunque x_0 appartiene all'insieme $f^{-1}(A)$ che per l'ipotesi è aperto su X : perciò $f^{-1}(A)$ è un intorno su X di x_0 .

Siccome:

$$x \in f^{-1}(A) \Rightarrow f(x) \in A \Rightarrow f(x) \in I'$$

si conclude che f è continua in x_0 . Per l'arbitrarietà di x_0 la a) è dimostrata.

b) $\Leftrightarrow$ c). Supponiamo vera la b). Se C è un insieme chiuso su $f(X)$, esiste un insieme A aperto su $f(X)$ tale che $C = f(X) - A$: per la prima delle (4) ne segue che $f^{-1}(C) = f^{-1}(f(X)) - f^{-1}(A)$; tale uguaglianza, avendo osservato dopo la (5) che si ha $f^{-1}(f(X)) = X$, diviene:

(6) $$f^{-1}(C) = X - f^{-1}(A) \, .$$

Per la b) l'insieme $f^{-1}(A)$ è aperto su X : allora per la (6) l'insieme $f^{-1}(C)$ è chiuso su X, e la c) è provata. Viceversa, nell'ipotesi che sia vera la c), basta ripetere il discorso appena concluso scambiando tra loro C ed A nonché le parole *chiuso* e *aperto*.

Per la dimostrazione del teorema nel caso alternativo dell'enunciato, vale il ragionamento precedente con S' in luogo di $f(X)$.

(7.5). *La funzione f sia definita in* $X \subseteq S$ *ed a valori in S'. Si ha:*
f continua ed x compatto $\Rightarrow$ $f(X)$ *compatto.*

DIM. Si tratta di dimostrare che ogni ricoprimento aperto $\mathcal{R}$ di $f(X)$ contiene un ricoprimento finito di $f(X)$.
Sinteticamente il ragionamento che esporremo si esprime come segue: per la continuità di f le controimmagini mediante f degli insiemi della classe $\mathcal{R}$ costituiscono un ricoprimento $\mathcal{S}$ di X aperto su X; essendo per ipotesi X compatto $\mathcal{S}$ contiene un ricoprimento finito di X, diciamolo $\mathcal{S}_0$: le immagini mediante f degli insiemi di $\mathcal{S}_0$ costituiscono un ricoprimento finito di $f(X)$ contenuto in $\mathcal{R}$, donde l'asserto.
Per la definizione di ricoprimento, $f(X)$ è contenuto nell'unione degli insiemi della classe $\mathcal{R}$:

$$f(X) \subseteq \bigcup_{B\in\mathcal{R}} B$$

e quindi per la seconda delle (4) si ha:

$$f^{-1}\big(f(X)\big) \subseteq f^{-1}\left(\bigcup_{B\in\mathcal{R}} B\right) \quad ,$$

cioè, essendo la controimmagine dell'unione uguale all'unione delle controimmagini:

$$f^{-1}\big(f(X)\big) \subseteq \bigcup_{B\in\mathcal{R}} f^{-1}(B) \ .$$

Per la prima delle (5) con $X^*=X$ ne segue :

$$X \subseteq \bigcup_{B\in\mathcal{R}} f^{-1}(B)$$

e quindi la famiglia $\mathcal{S} = \big(f^{-1}(B)\big)_{B\in\mathcal{R}}$ è un ricoprimento di X, che per la (7.4) di pag.22 [precisamente a) $\Leftrightarrow$ b)] è aperto su X in quanto B è aperto . Per l'ipotesi X compatto, e per la (3.3) di pag.14, $\mathcal{S}$ contiene un ricoprimento finito di X, sia esso:

$$\mathcal{S}_0 = \{ f^{-1}(B_1), \ldots, f^{-1}(B_k) \}.$$

Per la definizione di ricoprimento, essendo l'unione delle controimmagini uguale alla controimmagine dell'unione, si ha allora:

$$X \subseteq f^{-1}\left(\bigcup_{r=1}^{k} B_r\right)$$

Passando alle immagini mediante f , per la seconda delle (5) si ottiene :

$$f(X) \subseteq \bigcup_{r=1}^{k} B_r \ .$$

Pertanto la ennupla $\{B_1, \ldots, B_k\}$ è un ricoprimento di X contenuto in $\mathcal{R}$, e ciò prova l'asserto.

(7.6). *La funzione f sia definita in* $X \subseteq S$ *ed a valori in* S' *. Si ha:*
f continua e X connesso $\Rightarrow$ $f(X)$ *connesso .*

DIM. Procedendo per assurdo ammettiamo che $f(X)$ non sia connesso, sicchè esiste una partizione di $f(X)$ costituita da due insiemi non vuoti e aperti su $f(X)$, diciamoli A e B. Essendo f continua, per la (7.4) di pag.22, gli insiemi $f^{-1}(A)$ e $f^{-1}(B)$, ovviamente non vuoti, sono aperti su X : siccome essi costituiscono una partizione di X, ne segue che X non è connesso in contrasto con l'ipotesi.

Ricordando che i sottoinsiemi compatti di **R** sono dotati di minimo e di massimo, e che i sottoinsiemi connessi di **R** sono gli intervalli, da (7.5),(7.6) per $S' = \mathbf{R}$ si deduce :

(7.7). *La funzione f sia definita in* $X \subseteq S$ *ed a valori reali. Si ha:*
f continua e X compatto $\Rightarrow$ *f dotata di minimo e di massimo*
f continua e X connesso $\Rightarrow$ *il codominio di f è un intervallo.*

La prima implicazione è il teorema di Weierstrass , la seconda il teorema di Bolzano.

8. Confronto di topologie.

Siano $\mathcal{T}_1$ e $\mathcal{T}_2$ topologie per l'insieme *S*. Se risulta:

(1) $$\mathcal{T}_1 \subseteq \mathcal{T}_2$$

si dice che la topologia $\mathcal{T}_1$ è **più debole** della topologia $\mathcal{T}_2$, oppure che la seconda è **più forte (** o **più fine)** della prima.

Denotando con $\mathcal{T}_0$ la topologia banale per *S* e con $\mathcal{T}^*$ la topologia discreta per *S* , se $\mathcal{T}$ è una qualunque topologia per *S* si ha ovviamente:

$$\mathcal{T}_0 \subseteq \mathcal{T} \subseteq \mathcal{T}^* \quad ,$$

e perciò *di tutte le topologie per S quella banale è la più debole e quella discreta è la più forte.*

D'ora in poi per intendere che un termine , ad es. *aperto, chiuso, intorno, chiusura*, ecc. , si riferisce allo spazio topologico $(S, \mathcal{T}_1)$ [allo spazio topologico $(S, \mathcal{T}_2)$] , adopereremo quel termine preceduto dal numero 1 [dal numero 2] ; ad esempio

1-*aperto*, 1-*chiuso*, 1-*intorno*, 1-*chiusura*, ecc

[2-*aperto*, 2-*chiuso*, 2-*intorno*, 2-*chiusura*, ecc] .

Ciò premesso, rileviamo che la (1) si può esprimere come segue:

(2) $$A \text{ è un insieme 1-aperto} \Rightarrow A \text{ è 2-aperto} \quad ,$$

e che si ha pure :

(3) $$C \text{ è un insieme 1-chiuso} \Rightarrow C \text{ è 2-chiuso}$$

(4) $$I \text{ è un 1-intorno di } x_0 \Rightarrow I \text{ è un 2-intorno di } x_0$$

DIM.

$$C \text{ 1-chiuso} \Rightarrow C = -A \text{ con } A \text{ 1-aperto} \Rightarrow$$
$$\Rightarrow C = -A \text{ con } A \text{ 2-aperto} \Rightarrow C \text{ 2-chiuso}$$

$$I \text{ 1-intorno di } x_0 \Leftrightarrow \exists A \text{ 1-aperto tale che } x_0 \in A \subseteq I \Rightarrow$$
$$\Rightarrow \exists A \text{ 2-aperto tale che } x_0 \in A \subseteq I \Leftrightarrow I \text{ 2-intorno di } x_0$$

Dalla (4) si deduce che:

(5) $$\text{il punto } x_0 \text{ è 1-interno ad } X \Rightarrow \text{il punto } x_0 \text{ è 2-interno ad } X$$

(6) $$\text{il punto } x_0 \text{ non è 1-aderente ad } X \Rightarrow \text{il punto } x_0 \text{ non è 2-aderente ad } X$$

DIM. Si ha:

$$\begin{aligned} &\textit{il punto } x_0 \textit{ è 1-interno ad } X \Rightarrow \exists \textit{ un 1- intorno di } x_0 \textit{ contenuto in } X \Rightarrow \\ &\Rightarrow \exists \textit{ un 2- intorno di } x_0 \textit{ contenuto in } X \Rightarrow \\ &\Rightarrow \textit{ il punto } x_0 \textit{ è 2-interno ad } X \end{aligned}$$

$$\begin{aligned} &\textit{il punto } x_0 \textit{ non è 1-aderente ad } X \Rightarrow \exists \textit{ un 1- intorno di } x_0 \textit{ disgiunto da } X \Rightarrow \\ &\Rightarrow \exists \textit{ un 2- intorno di } x_0 \textit{ disgiunto da } X \Rightarrow \\ &\Rightarrow \textit{ il punto } x_0 \textit{ non è 2-aderente ad } X \end{aligned}$$

Notiamo che per la (5) si ha:

(7) $$1\text{-}\textit{interno di } X \subseteq \ 2\text{-}\textit{interno di } X \,,$$

e che per la (6) risulta $-(\ 1\text{-}\textit{chiusura di } X\ ) \subseteq\ -(\ 2\text{-}\textit{chiusura di } X\ )$, ossia:

(8) $$(\ 1\text{-}\textit{chiusura di } X\ ) \supseteq (\ 2\text{-}\textit{chiusura di } X\ ).$$

Le (7) e (8) si possono esprimere dicendo che se al sostegno di uno spazio topologico si associa un'altra topologia , più debole, l'interno di un insieme con la topologia più debole è contenuto nell'interno dello stesso insieme con la topologia più forte, mentre la chiusura di un insieme con la topologia più debole contiene la chiusura dello stesso insieme con la topologia più forte.

A differenza delle (5) e (6), per la convergenza di una successione non vale $1 \Rightarrow 2$, ma sussiste l'implicazione opposta; ciò accade anche per la compattezza di un insieme.

Precisamente si ha :

(9) $$\left(x_n\right)_{n\in N}\ 2\text{-}\textit{converge ad } x_0 \ \Rightarrow\ \left(x_n\right)_{n\in N}\ 1\text{-}\textit{converge ad } x_0 \ ,$$

(10) $$X \textit{ è un insieme } 2\text{-}\textit{compatto} \Rightarrow X \textit{ è un insieme } 1\text{-}\textit{compatto}$$

DIM. Occorre osservare che a causa della (4) la classe dei 2-intorni di x_0 contiene quella degli 1-intorni di x_0 , sicchè una proprietà vera per ogni 2-intorno è vera anche per ogni 1-intorno; analogamente, a causa della (2) la classe dei ricoprimenti 2-aperti di X contiene quella dei ricoprimenti 1-aperti di X , sicchè una proprietà vera per ogni ricoprimento 2-aperto è vera anche per ogni ricoprimento 1-aperto. Si ha dunque:

$$\begin{aligned} &\left(x_n\right)_{n\in N}\ \ 2\text{-}\textit{converge ad } x_0 \Rightarrow \\ &\forall\, I\ \ 2\text{-}\textit{intorno di}\ \ x_0\ \exists\ \nu \in \mathbf{N} : (\ n > \nu \Rightarrow x_n \in I\ ) \ \Rightarrow \\ &\Rightarrow \forall\, I\ \ 1\text{-}\textit{intorno di}\ \ x_0\ \exists\ \nu \in \mathbf{N} : (\ n > \nu \Rightarrow x_n \in I\ ) \ \Rightarrow \\ &\Rightarrow \left(x_n\right)_{n\in N}\ \ 1\text{-}\textit{converge ad } x_0 \ . \end{aligned}$$

$$\begin{aligned} &X\ \ 2\text{-}\textit{compatto} \Leftrightarrow \textit{Ogni ricoprimento } 2\text{-}\textit{aperto di } X \textit{ contiene un ricoprimento finito di } X \Rightarrow \\ &\Rightarrow \textit{Ogni ricoprimento } 1\text{-}\textit{aperto di } X \textit{ contiene un ricoprimento finito di } X \Leftrightarrow \\ &\Leftrightarrow\ X\ \ 1\text{-}\textit{compatto} \end{aligned}$$

Le (9), (10) si possono esprimere dicendo che se al sostegno di uno spazio topologico si associa un'altra topologia più debole, ogni successione convergente con la topologia più forte converge anche con quella più debole; analogamente, un insieme compatto con la topologia più forte è compatto anche con quella più debole.

OSSERVAZIONE

Dalle dimostrazioni di (5), (6) e da quelle di (9),(10) , emerge che si ha $1 \Rightarrow 2$ $[2 \Rightarrow 1]$ quando il significato del termine in questione si esprime con una proposizione che comincia con $\exists$ [con $\forall$] ; lo stesso accade quando il termine in questione è preceduto da $\exists$ [da $\forall$] . Nei casi come (7),(8) , in cui figura un'inclusione, si può utilizzare quanto sopra pensando ai punti dell'insieme e al significato dell'inclusione.

L'osservazione ci sembra utile per ricordare con facilità le precedenti asserzioni.

Dati gli spazi topologici S ed S', consideriamo ora una funzione continua f di S in S'.

In base alla precedente osservazione si verifica subito che : la proprietà di continuità di f si conserva se alla topologia esistente in S si sostituisce una topologia più forte, e/o alla topologia esistente in S' si sostituisce una topologia più debole.

* * *

Consideriamo una famiglia di topologie per S:

$$(\mathcal{T}_i)_{i \in I} \quad .$$

E' immediato constatare che l'insieme:

$$\mathcal{T} = \bigcap_{i \in I} \mathcal{T}_i$$

è a sua volta una topologia per S , che ovviamente è più debole di tutte le topologie della famiglia.

Scelta ad arbitrio una classe $\mathcal{A}$ di sottoinsiemi di S, non è detto che $\mathcal{A}$ sia una topologia per S; peraltro esiste ovviamente almeno una topologia per S alla quale appartengano tutti i sottoinsiemi della classe $\mathcal{A}$, ossia una topologia per S contenente $\mathcal{A}$: basta pensare alla topologia discreta.

Orbene, l'intersezione di tutte le topologie per S contenenti $\mathcal{A}$, che è una topologia per S, è *la più debole fra le topologie per S contenenti* $\mathcal{A}$. Per così dire, è "la migliore" topologia da associare ad S in modo che fra gli aperti dello spazio topologico vi siano tutti i sottoinsiemi della classe $\mathcal{A}$ prescelta.

ESEMPI ED OSSERVAZIONI .

1. Detto S un insieme non vuoto, costituito ad almeno due elementi, si associ ad S la topologia discreta $\mathcal{T} = \mathcal{P}(S)$.

Nello spazio topologico $(S, \mathcal{T})$ si ha quanto segue:

- Tutti i sottoinsiemi di S sono simultaneamente aperti e chiusi, quindi per qualunque sottoinsieme X di S l'interno e la chiusura di X coincidono entrambi con X.
- Lo spazio non è connesso per quanto sopra .
- Per ogni $x_0 \in S$, un intorno di x_0 è un qualunque insieme al quale appartenga x_0 . In particolare, l'insieme $\{x_0\}$ è un intorno di x_0 .
- Per qualunque sottoinsieme X di S risulta $\partial X = \emptyset$ e $Dr(X) = \emptyset$.
- $(S, \mathcal{T})$ è uno spazio di Hausdorff ([1]).
- Le successioni convergenti sono solo quelle costanti ([2]).
- I soli sottoinsiemi compatti sono quelli finiti ([3]).
- I soli sottoinsiemi connessi sono quelli costituiti da un unico elemento ([4]).

2. Detto S un insieme non vuoto, costituito da almeno due elementi, si associ ad S la topologia banale $\mathcal{T} = \{\emptyset, S\}$.

Nello spazio topologico $(S, \mathcal{T})$ si ha quanto segue:

- I soli sottoinsiemi aperti e chiusi sono $\emptyset$ ed S; per qualunque insieme diverso da S l'interno è vuoto; per qualunque insieme non vuoto la chiusura è S.
- Ogni punto ha un solo intorno, che è l'insieme S.
- Qualunque sia X non vuoto e diverso da S risulta $\partial X = S$ e $Dr(X) = S$.
- $(S, \mathcal{T})$ non è uno spazio di Hausdorff , ed è uno spazio connesso.
- Tutte le successioni sono convergenti, ed hanno per limite qualunque punto.
- Ogni insieme è compatto; ogni insieme non vuoto è connesso ([5]).

([1]) Se x ed y sono punti distinti di S, gli intorni $\{x\}$ di x ed $\{y\}$ di y sono disgiunti.

([2]) Se la successione $(x_n)_{n \in N}$ converge ad x_0, scelto l'intorno $\{x_0\}$ di x_0 esiste $\nu \in \mathbf{N}$ tale che per $n > \nu$ si abbia $x_n \in \{x_0\}$, ossia $x_n = x_0$.

([3]) Se $X = \{x_1, x_2, \ldots, x_k\}$ e $\mathcal{R} = (A_i)_{i \in I}$ è un ricoprimento aperto di X, esistono $i_1 \in I$ tale che $x_{i_1} \in A_{i_1}$, … , $i_k \in I$ tale che $x_{i_k} \in A_{i_k}$: allora $\{A_{i_1}, ..., A_{i_k}\}$ è un ricoprimento finito di X contenuto in $\mathcal{R}$. Pertanto ogni insieme finito è compatto.

Viceversa, sia X un insieme compatto. La famiglia $(\{x\})_{x \in X}$ è un ricoprimento aperto di X , e quindi deve contenere un ricoprimento finito di X : dunque X è incluso nell'unione di un numero finito di insiemi costituiti da un solo elemento , e pertanto è un insieme finito.

([4]) Sia X un insieme costituito da almeno due elementi. Scelto un punto $x_0 \in X$, gli insiemi non vuoti $\{x_0\}$, $X - \{x_0\}$ costituiscono una partizione di X e sono aperti su X : pertanto X non è connesso.

([5]) Siano X un sottoinsieme di S ed $\mathcal{R}$ un ricoprimento aperto di X. Siccome gli aperti sono soltanto $\emptyset$ ed S, e gli insiemi aperti della famiglia $\mathcal{R}$ non possono essere tutti uguali a $\emptyset$, almeno uno si essi è S : dunque $\mathcal{R}$ contiene il ricoprimento finito di X costituito dal solo elemento S, e pertanto X è compatto. Se $X \neq \emptyset$, siccome l'unico aperto non vuoto è S non esiste alcuna partizione di X costituita da due aperti non vuoti : pertanto X è connesso.

3. Sia S un insieme costituito da due elementi, diciamoli a e b. E' facile verificare che la classe $\mathcal{T}$ di sottoinsiemi di S costituita da S, $\varnothing$, $\{b\}$ è una topologia per S.

Nello spazio topologico $(S, \mathcal{T})$ si ha quanto segue:

3_1 – Gli insiemi chiusi sono $\varnothing$, S, $\{a\}$.

3_2 – Lo spazio è connesso ([1]).

3_3 – L'interno dell'insieme $\{a\}$ è vuoto, la chiusura è $\{a\}$ ([2]).

3_4 – L'interno dell'insieme $\{b\}$ è $\{b\}$, la chiusura è S.

3_5 – L'insieme $\{b\}$ è denso in S [pag.19 e 3_4].

3_6 – Del punto a esiste un solo intorno, che è S ; gli intorni di b sono $\{b\}$ ed S.

3_7 – Per la frontiera e il derivato degli insiemi $\{a\}$ e $\{b\}$ si ha:

$$\partial\{a\} = \{a\} \ , \ \partial\{b\} = \{a\} \ \text{ e } \ Dr(\{a\}) = \varnothing \ , \ Dr(\{b\}) = \{a\} \ . \quad ([3])$$

3_8 – $(S, \mathcal{T})$ non è uno spazio di Hausdorff ([4]).

3_9 – Una successione $\{x_n\}$ converge al punto b se e solo se risulta definitivamente $x_n = b$, sicché le successioni convergenti a b sono quelle definitivamente costanti, con la costante uguale a b ([5]).

3_{10} – Tutte le successioni (anche quelle convergenti a b) convergono al punto a ([6]).

Il lettore osservi che nello spazio topologico **R** *con la topologia usuale* [cfr.inizio pag.9] *le precedenti asserzioni* (con a e b numeri reali), *esclusa la seconda e la terza, sono false*.

La 3_4 è falsa perché, con riferimento ad un insieme costituito da un solo elemento, l'interno è vuoto e la chiusura è l'insieme stesso ; la 3_7 è falsa per l'elemento b, dato che un insieme costituito da un solo elemento ha per frontiera l'insieme stesso e per derivato l'insieme vuoto; la 3_9 è falsa, perché ad esempio la successione $\{1/n\}$ converge a 0 ma non è definitivamente costante. Che siano false le rimanenti (escluse seconda e terza) è banale.

([1]) Per la (3.2) di pag.13 [equivalenza tra a) e c)] basta osservare che gli insiemi simultaneamente aperti e chiusi sono soltanto $\varnothing$ ed S.

([2]) L'interno di $\{a\}$ è il più grande degli aperti (ossia degli elementi di $\mathcal{T}$) contenuti in $\{a\}$ [(1.2) pag.6] : ricordando la definizione di $\mathcal{T}$, ovviamente è $\varnothing$. La chiusura di $\{a\}$ è il più piccolo dei chiusi contenenti $\{a\}$, che è $\{a\}$ [cfr. 3_1].

([3]) Tenendo presente 3_6 si osserva quanto segue. All'unico intorno di a (che è S) appartengono l'unico punto dell'insieme $\{a\}$ [dell'insieme $\{b\}$] e l'unico punto del complementare (ossia di $\{b\}$ [di $\{a\}$]) : dunque il punto a è di frontiera per entrambi gli insiemi$\{a\}$ e $\{b\}$; invece il punto b non è di frontiera per alcuno dei due : infatti esiste un intorno di b, precisamente $\{b\}$, a cui non appartiene il punto a , che è l'unico punto del primo insieme e l'unico punto del complementare del secondo.

Le precedenti osservazioni provano le prime due uguaglianze. Per la terza è immediato constatare che i punti a e b non sono di accumulazione per l'insieme $\{a\}$. Quanto alla quarta, basta osservare che non esistono punti dell'insieme $\{b\}$ diversi da b (sicché b non è di accumulazione per l'insieme), mentre all'unico intorno del punto a (cioè ad S) appartiene b che è un punto dell'insieme diverso da a (sicché a è di accumulazione per l'insieme).

([4]) Due punti distinti di S sono soltanto a e b : non esistono un intorno di a e un intorno di b disgiunti perché l'unico intorno di a è S [per 3_6] .

([5]) Infatti, scelto l'intorno $\{b\}$ del punto b [cfr. 3_6] , esiste $\nu \in \mathbf{N}$: $n > \nu \Rightarrow x_n \in \{b\}$ ossia $x_n = b$.

([6]) Con riferimento ad una successione qualunque, basta osservare che all'unico intorno del punto a (cioè S, per3_6) appartengono tutti i termini della successione.

4. Ricordiamo che un insieme si dice **discreto** quando o è finito o è numerabile.

Diciamo X un sottoinsieme di **R** avente la seguente proprietà :

$-X$, cioè il complementare di X, è un insieme discreto .

Si verifica facilmente che la classe $\mathcal{T}$ costituita da $\varnothing$, **R** e dai sottoinsiemi di **R** aventi la proprietà del predetto X , è una topologia per **R**.

Nello spazio topologico (**R**, $\mathcal{T}$) si ha quanto segue:

4_1 − Gli aperti sono $\varnothing$, **R** e i sottoinsiemi di **R** aventi il complementare discreto.

4_2 − I chiusi sono **R** , $\varnothing$ e gli insiemi discreti .

4_3 − Qualunque insieme che contiene un aperto non vuoto è a sua volta aperto ([7]).

4_4 − Gli intorni di un dato punto x_0 sono gli aperti a cui x_0 appartiene.

4_5 − Qualunque insieme infinito non numerabile è denso in **R** [pag.19] ([8]).

4_6 − Due aperti qualunque, non vuoti, hanno almeno un punto in comune ([9]) .

4_7 − (**R**, $\mathcal{T}$) non è uno spazio di Hausdorff ([10]).

4_8 − L'insieme **R** è connesso ([11])

4_9 − Gli insiemi non connessi sono quelli discreti aventi almeno due elementi ([12])

4_{10} − Le successioni convergenti sono quelle definitivamente costanti ([13])

4_{11} − Ogni successione convergente ha un unico limite ([14]).

Le dimostrazioni di 4_8, …. , 4_{11} , esposte nelle note ([11]) , … , ([14]), sono riportate nella pagina seguente.

([7]) Sia B un insieme che contiene un aperto $A \neq \varnothing$. Se $A = \mathbf{R}$, B è uguale ad **R** e quindi è aperto.
Supposto $A \neq \mathbf{R}$ l'insieme chiuso $-A$, essendo diverso da $\varnothing$ e da **R**, è discreto [per 4_2]. Allora :

$$B \supseteq A \Rightarrow -B \subseteq -A \Rightarrow -B \text{ discreto} \Rightarrow -B \text{ chiuso } [\text{per } 4_2] \Rightarrow B \text{ aperto} .$$

([8]) Sia X un insieme infinito non numerabile. L'insieme chiusura di X non è finito (perché contiene [(1) pag.6] l'insieme infinito X) e non è numerabile (se fosse numerabile tale sarebbe la sua parte infinita X), quindi non è discreto (per la definizione di discreto); inoltre, banalmente non è $\varnothing$. Pertanto la chiusura di X, essendo un insieme chiuso [(1.3) pag.6] , per 4_2 è uguale ad **R**.

([9]) Siano A_1, A_2 aperti non vuoti . Poichè l'asserto è banale se uno dei due insiemi è **R,** supponiamo che essi siano entrambi diversi da **R** , oltre che non vuoti : pertanto i rispettivi insiemi complementari sono entrambi diversi da $\varnothing$ e da **R** . Orbene, se fosse $A_1 \cap A_2 = \varnothing$, sarebbe $-(A_1 \cap A_2) = \mathbf{R}$, ossia [cfr.([1]) pag.5] $(-A_1) \cup (-A_2) = \mathbf{R}$, il che è assurdo in quanto gli insiemi chiusi al primo membro sono discreti (per 4_2 , dato che sono diversi da **R** e da $\varnothing$) **.**

([10]) Un intorno qualunque di un punto qualunque, poiché contiene un aperto non vuoto, per 4_3 è un insieme aperto. In conseguenza, per 4_6, due intorni qualunque di due punti distinti hanno almeno un punto in comune.

DIM. 4_8 ([11]) DIM. 4_9 ([12]) DIM. 4_{10} ([13]) DIM. 4_{11} ([14])

E' utile evidenziare che nello spazio in questione *sussiste l'unicità del limite di una successione convergente pur non essendo uno spazio di Hausdorff* [cfr. 4_7, 4_{11}].

Il lettore osservi che nello spazio topologico **R** *con la topologia usuale le precedenti asserzioni, escluse le* 4_8 , 4_{11} *, sono false.* Qui di seguito, accanto al simbolo che denota ciascuna di esse, riportiamo una motivazione a prova che è falsa (qualora si intenda nello spazio topologico **R** con la topologia usuale).

4_1 − L'intervallo]-1,1[è un aperto, non è ∅, **R** , e il complementare non è discreto.
4_2 − L'intervallo [-1,1] è un chiuso, non è ∅, **R** e non è discreto.
4_3 − L'intervallo [-1,1] contiene infiniti aperti non vuoti, ma non è aperto.
4_4 − L'intervallo [-1,1] è un intorno di 0, ma non è un aperto a cui 0 appartiene.
4_5 − L'intervallo [-1,1] è un insieme infinito non numerabile, ma non è denso in **R**.
4_6 − Gli intervalli]-1,1[,]2,3[sono due aperti, ma non hanno punti in comune.
4_7 − Lo spazio è uno spazio di Hausdorff.
4_9 − L'insieme $[-1,1]\cup\{2\}$ non è connesso, ma non è un insieme discreto.
4_{10} − La successione $\{1/n\}$ converge a 0 ma non è definitivamente costante.

([11]) Procedendo per assurdo, ammettiamo che **R** non sia connesso. Per la (3.2) di pag.13 [equivalenza tra a) e c)] esiste allora una partizione di **R** costituita da due insiemi chiusi e non vuoti. Tali insiemi chiusi sono quindi entrambi diversi da **R** e da ∅, sicché per 4_2 sono entrambi discreti : ma ciò è assurdo perché l'unione di essi, ovviamente insieme discreto, per la definizione di partizione è uguale ad **R** (che non è discreto).

([12]) Sia X un sottoinsieme di **R** non connesso: per la (3.1) di pag.12 esiste una partizione di X , diciamola $\{X_1, X_2\}$, costituita da due insiemi non vuoti e chiusi su X ; esistono allora [pag.9] due chiusi C_1, C_2 tali che $X_1=X\cap C_1$, $X_2=X\cap C_2$. Risulta $C_1\neq$ **R** : se fosse C_1=**R** sarebbe $X_1=X$ e quindi, intersecando con X_2 ambo i membri di tale uguaglianza si avrebbe $X_1 \cap X_2 = X \cap X_2$ ossia $\varnothing = X_2$, il che è assurdo; inoltre risulta $C_1\neq\varnothing$: se fosse $C_1 = \varnothing$ sarebbe $X_1=\varnothing$, il che è assurdo. Da quanto stabilito si deduce che l'insieme chiuso C_1 è discreto [per 4_2], ed allora anche $X_1\subseteq C_1$ è discreto. Analogamente si prova che X_2 è discreto, e quindi $X = X_1 \cup X_2$ è un insieme discreto. Che X ha almeno due elementi è evidente, dato che X è l'unione degli insiemi X_1, X_2 disgiunti e non vuoti.

Resta da dimostrare che viceversa un insieme X discreto, avente almeno due elementi, non è connesso.

Scelto un punto $a\in X$, gli insiemi $\{a\}$, X - $\{a\}$ sono non vuoti (il secondo perché X ha almeno due elementi) e sono entrambi chiusi (per 4_2) : siccome essi costituiscono una partizione dell'insieme X chiuso (in quanto discreto), si conclude che X non è connesso [pag.12 ultimi tre righi.]

([13]) Consideriamo una successione $\{x_n\}$ di numeri reali convergente ad un punto x_0, e denotiamo con X l'insieme costituito dai termini della successione diversi da x_0 . Essendo X discreto, quindi chiuso [per 4_2], l'insieme $-X$ è un aperto al quale x_0 appartiene e perciò è un intorno di x_0 [cfr. 4_4] : allora esiste $\nu\in$ **N** tale che per $n>\nu$ risulta $x_n\in -X$. Dunque per $n>\nu$ il punto x_n non è diverso da x_0 , ossia è uguale ad x_0 : pertanto la successione è definitivamente costante.

([14]) Detta $\{x_n\}$ una successione convergente sia ad a che a b , ripetendo il ragionamento della nota ([13]) prima con $x_0=a$ e poi con $x_0=b$, si prova che esistono $\nu,\nu_1\in$ **N** tali che: $x_n=a\ \forall\, n>\nu$, $x_n=b\ \forall\, n>\nu_1$. Scelto un $m\in$ **N** maggiore sia di ν che di ν_1 , si ha $x_m=a$, $x_m=b$, donde $a=b$.

5. Consideriamo i sottoinsiemi di R ai quali non appartiene 1, nonché quelli ai quali appartiene 1 che contengano almeno un intervallo del tipo $]0,1/k[$ con $k \in N$.

Si verifica facilmente che la totalità $\mathcal{C}^*$ dei sottoinsiemi di R dianzi considerati è una topologia per R, sicché la coppia $(R, \mathcal{C}^*)$ è uno spazio topologico. Con riferimento a tale spazio si ha ad esempio che:

— I seguenti insiemi sono aperti perché 1 non vi appartiene:

(1) $]-3,0]$, $]-\infty,1[$, $[0,1[$, $]0,1[$, $[2,3]$, $\{2\}$, $[0,2]-\{1\}$, $Q \cap]0,\frac{1}{2}]$

— Anche i seguenti insiemi sono aperti, perché 1 vi appartiene e ciascuno di essi contiene un intervallo del tipo $]0,1/k[$ con $k \in N$:

(2) $]0,\frac{1}{3}] \cup \{1\}$, $[0,2]$, $]-\infty,2]$, $[0,+\infty[$, $\{x \in Q : x < -2\} \cup [0,3]$

Notiamo che gli insiemi (2) sono anche chiusi, perché i rispettivi insiemi complementari (ai quali non appartiene 1) sono aperti; invece degli insiemi aperti (1) solo alcuni sono anche chiusi, precisamente $]-3,0]$, $[2,3]$, $\{2\}$, perché i rispettivi complementari (ai quali appartiene 1) sono aperti.

— I seguenti insiemi sono chiusi in quanto vi appartiene 1 (sicché 1 non appartiene ai complementari e perciò questi sono aperti):

(3) $]-\infty,1]$, $[1,+\infty[$, $\{1\}$, $R-\{0\}$, $(R-[0,2]) \cup \{1\}$, $[0,1] \cap Q$

— Anche i seguenti insiemi sono chiusi, perché i complementari sono aperti:

(4) $]-\infty,0] \cup]\frac{3}{2},+\infty[$, $]2,+\infty[$, $]2,4[$, $[5,+\infty[$.

Notiamo che gli insiemi (4) sono anche aperti, perché non vi appartiene 1; invece degli insiemi chiusi (3) solo alcuni sono anche aperti, precisamente $]-\infty,1]$ e $R-\{0\}$.

Per lo spazio topologico $(R, \mathcal{C}^*)$ si ha quanto segue.

— Gli intorni del punto 1 sono gli insiemi del tipo

(5) $\qquad I = X \cup]0, 1/k[\cup \{1\} \qquad$ con $X \subseteq R$ e $k \in N$ [1].

Per ogni punto $x_0 \neq 1$, un intorno di x_0 è qualunque insieme al quale appartenga x_0 [2].

— Esistono insiemi simultaneamente aperti e chiusi (come si è visto), oltre a R e a $\emptyset$: perciò lo spazio non è connesso [Teor. (3.2) pag. 13].

— $(R, \mathcal{T}^*)$ è uno spazio di Hausdorff [3]

— Qualunque sia la successione $\{x_n\}$ di numeri reali convergente, si ha:

$\lim x_n = l \neq 1 \implies$ la successione $\{x_n\}$ è definitivamente costante [4].

— Detta $\mathcal{T}$ la topologia usuale per R, sia $\{x_n\}$ una qualunque successione convergente a 0 nello spazio topologico $(R, \mathcal{T})$ e definitivamente positiva. Allora nello spazio topologico $(R, \mathcal{T}^*)$ si ha:

$$\lim x_n = 1 \qquad [5].$$

---

(1) Infatti un insieme del tipo (5) contiene $]0, 1/k[\cup \{1\}$, che è un aperto al quale appartiene 1.

(2) Un insieme siffatto contiene $\{x_0\}$, e questi è un aperto (dato che 1 non vi appartiene) al quale appartiene x_0.

(3) Siano x, y punti distinti di R. Se entrambi sono diversi da 1, gli insiemi disgiunti $\{x\}, \{y\}$ sono rispettivamente un intorno di x e un intorno di y. Esaminiamo ora il caso che uno dei due punti, ad es. y, sia 1. Se $x \leq 0$ oppure $x > 1$, gli insiemi $\{x\}$ e $]0, \frac{1}{2}[\cup \{1\}$ sono intorni disgiunti rispettivamente di x ed 1; se $0 < x < 1$, gli insiemi $\{x\}$ e $]0, 1/k[\cup \{1\}$ con $k \in N, \frac{1}{k} < x$ sono intorni disgiunti rispettivamente di x ed 1.

(4) Scelto come intorno di l l'insieme $\{l\}$, deve esistere $\nu \in N$ tale che per $n > \nu$ si abbia $x_n \in \{l\}$, ossia $x_n = l$.

(5) Un qualunque intorno di 1 è del tipo (5) con $X \subseteq R$ e $k \in N$. Siccome $x_n \to 0$ e $x_n > 0$ definitivamente, esiste $\nu \in N$ tale che per $n > \nu$ si abbia $0 < x_n < 1/k$ e quindi anche $x_n \in I$.

In particolare, nello spazio topologico $(R,\mathcal{T}^*)$ si ha:

$$\lim \frac{1}{n} = 1 \quad , \quad \lim \frac{3}{n} = 1 \quad , \quad \lim \frac{4}{n} = 1 \ .$$

Di conseguenza il limite della somma non è uguale alla somma dei limiti.

Sia $\{x_n\}$ una successione convergente a 0 in $(R,\mathcal{T})$, che abbia infiniti termini negativi. Allora la successione non ammette limite in $(R,\mathcal{T}^*)$ (1). In particolare il limite della successione $\{-1/n\}$ in $(R,\mathcal{T}^*)$ non esiste.

— Nello spazio topologico $(R,\mathcal{T}^*)$ gli insiemi compatti ai quali non appartenga 1 sono necessariamente finiti (2).

— La funzione f definita ponendo:

$$f(x) = (x-3)^2 \quad \forall x \in R,$$

come funzione di $(R,\mathcal{T})$ in $(R,\mathcal{T}^*)$ non è continua, in quanto risulta:

$$\lim_{x\to 3} f(x) = 1 \ . \tag{6}$$

Il lettore verifichi che per $x_0 \neq 3$ il limite di $f(x)$ per $x \to x_0$ non esiste.

6. Topologia della convergenza puntuale.

Sia S un insieme di funzioni reali definite in un sottoinsieme

---

(1) Sia per assurdo $\lim x_n = l$. Non può essere $l=1$, in quanto per infiniti n si ha $x_n < 0$ e quindi per infiniti n il termine x_n non appartiene all'intorno $]0,1/n[\cup \{1\}$ del punto 1; non può essere $l \neq 1$, altrimenti la successione sarebbe definitivamente costante (come già osservato) e quindi sarebbe identicamente nulla, in contrasto con l'ipotesi.

(2) Sia X un insieme compatto al quale non appartenga 1. Siccome la famiglia $(\{x\})_{x\in X}$ è un ricoprimento aperto di X, essa deve contenere un ricoprimento finito di X: esistono allora $x_1,\dots,x_n \in X$ tali che $X \subseteq \bigcup_{i=1}^{n} \{x_i\}$. Pertanto $X = \{x_1, x_2, \dots, x_n\}$.

(3) Sia I un qualunque intorno di 1 in $(R,\mathcal{T}^*)$, quindi un insieme del tipo (5). Scelto $\delta > 0$ tale che $\delta < 1/\sqrt{2}$, per $|x-3| < \delta$ e $x \neq 3$ si ha $(x-3)^2 \in I$, e ciò prova la (6).

A di R^n. Fissato un elemento di S, diciamolo $x = x(t) = x(t_1, \dots, t_n)$, scegliamo un numero positivo ε ed un sottoinsieme finito A^* di A; consideriamo quindi la parte di S:

$$I_x(\varepsilon, A^*) = \{ y \in S : |y(t) - x(t)| < \varepsilon \ \ \forall t \in A^* \}.$$

Al variare di ε in R^+ e di A^* fra i sottoinsiemi finiti di A, $I_x(\varepsilon, A^*)$ descrive un insieme di parti di S, diciamolo $I(x)$; aggreghiamo poi ad $I(x)$ tutte le parti di S contenenti almeno un elemento di $I(x)$. In tal modo si ottiene un insieme di parti di S che dipende dal fissato $x \in S$: diciamo $\mathcal{J}(x)$ tale insieme.

E' facile verificare che la famiglia $(\mathcal{J}(x))_{x \in S}$ soddisfa agli assiomi di Hausdorff, e quindi esiste una ben determinata topologia per S, diciamola $\mathcal{T}$, tale che per ogni $x \in S$ gli intorni di x sono gli elementi della famiglia $\mathcal{J}(x)$ [(4.1) pag. 16].

Dimostriamo che:

Lo spazio topologico $(S, \mathcal{T})$ è uno spazio di Hausdorff.

DIM. Fissati $x_1, x_2 \in S$ con $x_1 \neq x_2$, esiste almeno un $t_0 \in A$ in cui le funzioni $x_1(t)$, $x_2(t)$ assumono valori diversi; scelto $\varepsilon > 0$ tale che $2\varepsilon < |x_1(t_0) - x_2(t_0)|$, consideriamo gli intorni $I_1 = I_{x_1}(\varepsilon, A^*)$, $I_2 = I_{x_2}(\varepsilon, A^*)$, con $A^* = \{t_0\}$, rispettivamente di x_1 ed x_2. Evidentemente I_1 ed I_2 sono disgiunti; infatti se esistesse $y \in I_1 \cap I_2$, si avrebbe:

$$|x_1(t_0) - x_2(t_0)| \leq |x_1(t_0) - y(t_0)| + |y(t_0) - x_2(t_0)| < 2\varepsilon$$

in contrasto con la scelta di ε.

Sia $\{x_n\}$ una successione di punti di S. E' facile verificare che per $l \in S$ le due proposizioni:

(7) $\lim x_n = l$ nello spazio topologico $(S, \mathcal{T})$

(8) $\lim x_n = l$ puntualmente in A, cioè $\lim x_n(t) = l(t) \ \ \forall t \in A$

sono equivalenti. Per tal motivo la $\mathcal{T}$ si chiama topologia della convergenza puntuale per l'insieme S.

Il lettore osservi che se l è una funzione definita in A che non appartenga al considerato insieme S, può accadere che la (8) sia vera ma per la predetta l la (7) non ha senso.

Capitolo II - Cenni sulla teoria della misura e dell'integrazione secondo Lebesgue . Cenni sull'integrale di Stieltjes.

Gli argomenti del presente capitolo saranno accennati sommariamente; i risultati, dei quali non riporteremo la dimostrazione, non figureranno sempre come enunciati di teoremi: talvolta essi saranno esposti in forma di asserzioni nel contesto del discorso . Ciò in quanto gli argomenti in questione, che potrebbero essere oggetto di un intero corso, si possono considerare estranei all'Analisi Funzionale, pur essendo indispensabili per le applicazioni.

1 . La misura secondo Lebesgue.

Supporremo R^n munito della topologia usuale, che si può ad esempio introdurre come l'unica topologia avente per base la totalità degli intervalli aperti di R^n : pertanto un sottoinsieme di R^n é aperto se e solo se esso é l'unione di una famiglia di intervalli aperti.

Peraltro ogni aperto limitato di R^n é l'unione di un'infinità numerabile di intervalli aperti a destra (1) a due a due disgiunti.

Qualunque sia l'intervallo I di R^n aperto a destra e non vuoto , denoteremo con $m(I)$ la misura elementare di I , ossia il prodotto delle dimensioni di I ; se $I = \emptyset$, si porrà $m(I) = 0$.

La definizione della misura secondo Lebesgue che ci proponiamo di accennare, é così schematizzabile:

a) Si definisce la misura degli aperti limitati .

b) Si definisce la misura dei chiusi limitati.

c) Si definisce la misurabilità (e quindi la misura) per gli insiemi limitati.

d) Si definisce la misurabilità (e quindi la misura) per gli insiemi non necessariamente limitati.

a) Qualunque sia A aperto limitato di R^n , consideriamo una successione $(I_n)_{n\in N}$ di intervalli aperti a destra, a due a due disgiunti, l'unione dei quali sia uguale ad A , e denotiamo con I un intervallo aperto a destra contenente A. Risultando:

$$\sum_{i=1}^{n} m(I_i) \leq m(I) \quad \forall n \in N ,$$

---

(1) In questi appunti è rimasta la locuzione arcaica intervallo "aperto a destra": è preferibile dire "semiaperto a destra", dato che non si tratta di un insieme aperto.

la serie di termine generale $m(I_n)$ è convergente. Il numero reale:

$$\sum_{i=1}^{\infty} m(I_i)$$

è indipendente dalla scelta della successione $(I_n)_{n\in N}$: esso si chiama misura secondo Lebesgue dell'insieme A, e sarà denotato con m(A).

b) Qualunque sia C chiuso limitato di R^n, consideriamo un aperto limitato A di R^n contenente C. Il numero reale $m(A)-m(A-C)$ [1] è indipendente dall'aperto A, e si chiama misura secondo Lebesgue dell'insieme C. Tale numero sarà denotato con $m(C)$:

$$m(C)=m(A)-m(A-C)\ .$$

c) Sia X un sottoinsieme limitato di R^n. Denotiamo con $\mathcal{C}_X$ l'insieme numerico descritto dalle misure secondo Lebesgue dei chiusi (limitati) contenuti in X, e con $\mathcal{A}_X$ l'insieme numerico descritto dalle misure secondo Lebesgue degli aperti limitati contenenti X. Gli insiemi numerici $\mathcal{C}_X$ e $\mathcal{A}_X$ sono separati. Se essi sono anche contigui, si dice che X è misurabile secondo Lebesgue, e si pone:

$$m(X)=\sup \mathcal{C}_X=\inf \mathcal{A}_X\ .$$

Gli aperti limitati e i chiusi limitati sono misurabili secondo Lebesgue, e la misura di essi, secondo la definizione ora introdotta, coincide con quella definita in a) e b). Analogamente dicasi per gli intervalli aperti a destra.

L'intersezione di due insiemi limitati e misurabili secondo Lebesgue è un insieme misurabile secondo Lebesgue.

Qualunque siano X_1, X_2 limitati e misurabili secondo Lebesgue, si ha che :

$$X_1 \subseteq X_2 \Rightarrow m(X_1)\leq m(X_2)\ .$$

[1] Si osservi che A-C è un aperto limitato, sicché ha senso $m(A-C)$.

d) Sia X un sottoinsieme di R^n (limitato o non). Si dice che X é misurabile secondo Lebesgue quando per ogni sottoinsieme Y, limitato e misurabile secondo Lebesgue, l'insieme limitato:

$$X \cap Y$$

é misurabile secondo Lebesgue.

Nel caso che X sia limitato , la nozione ora introdotta é equivalente a quella introdotta in c).

Supposto X misurabile secondo Lebesgue (X limitato o non),sia $M(X)$ l'insieme numerico descritto da $m(X \cap Y)$ al variare di Y tra gli insiemi limitati e misurabili. L'elemento di $\overline{R}$ [1]:

$$\sup M(X)$$

si chiama misura di X secondo Lebesgue , e sarà denotato con $m(X)$.

Nel caso che X sia limitato, la misura di X secondo la definizione ora introdotta, coincide con quella definita in c).

Denotato con $\mathcal{L}$ l'insieme delle parti di R^n (limitate o non) misurabili secondo Lebesgue,l'applicazione:

$$m : X \in \mathcal{L} \longrightarrow m(X) \in [0, +\infty]$$

si chiama la misura di Lebesgue.

L'insieme $\mathcal{L}$ gode della seguente proprietà : ogni unione finita o numerabile di elementi di $\mathcal{L}$ appartiene ad $\mathcal{L}$; ogni intersezione finita o numerabile di elementi di $\mathcal{L}$ appartiene ad $\mathcal{L}$; la differenza di due elementi di $\mathcal{L}$ appartiene ad $\mathcal{L}$. Ogni insieme numerabile é misurabile secondo Lebesgue, ed ha misura nulla (ciò consegue dalla "numerabile additività" formulata più avanti).

Tutte le proprietà della misura di Peano-Jordan sono valide anche per la misura di Lebesgue. In più, per ogni successione $(X_n)_{n \in N}$ di elementi di $\mathcal{L}$ a due a due disgiunti , risulta;

$$m\left(\bigcup_{n \in N} X_n\right) = \sum_{n \in N} m(X_n) \ ;$$

---

(1) Con $\overline{R}$ denotiamo l'insieme ampliato dei numeri reali.

tale proprietà si esprime dicendo che la misura di Lebesgue è numerabilmente additiva.

Ogni insieme misurabile secondo Peano-Jordan è altresì misurabile secondo Lebesgue, e la sua misura secondo Peano-Jordan coincide con quella secondo Lebesgue: in altri termini, detto $\mathcal{M}$ l'insieme delle parti di R^n misurabili secondo Peano-Jordan, risulta $\mathcal{M} \subset \mathcal{L}$ e la misura di Lebesgue è un prolungamento su $\mathcal{L}$ della misura di Peano-Jordan; anzi, è l'unico prolungamento su $\mathcal{L}$ di detta misura che gode della proprietà di numerabile additività.

E' da rilevare che, a differenza di quanto accade per gli elementi di $\mathcal{M}$, gli insiemi misurabili secondo Lebesgue non hanno necessariamente frontiera di misura nulla.

D'ora in avanti, salvo esplicito avviso contrario, la parola "misurabile" sarà intesa nel senso di Lebesgue, e con la lettera m si denoterà sempre la misura secondo Lebesgue.

2. Funzioni misurabili.

Sia f una funzione reale definita nel sottoinsieme X di R^n; con X misurabile.

Si dice che f è misurabile quando per ogni $a \in R$ l'insieme:

(1) $$\{x \in X : f(x) > a\}$$

è misurabile (1).

La somma, la differenza, il prodotto, il rapporto (se non si annulla la funzione a denominatore) di funzioni misurabili, sono funzioni misurabili. Se f è misurabile, tali sono $\lambda f \ \forall \lambda \in R$ e $|f|^\lambda \ \forall \lambda \in R^+$.

Il limite di una successione di funzioni misurabili convergente puntualmente, è una funzione misurabile.

Per le funzioni f limitate, definite in un sottoinsieme X di R^n misurabile

---

(1) Tale circostanza equivale a ciascuna delle seguenti:

1) $\forall a \in R$, l'insieme $\{x \in X : f(x) \geq a\}$ è misurabile.
2) $\forall a \in R$, l'insieme $\{x \in X : f(x) < a\}$ è misurabile.
3) $\forall a \in R$, l'insieme $\{x \in X : f(x) \leq a\}$ è misurabile.

secondo Peano-Jordan e di misura finita, l'integrabilità secondo Riemann (1) si caratterizza in termini di continuità, in quanto (criterio di Vitali-Lebesgue):

$$(*) \qquad \left(\begin{matrix} f \text{ integrabile in } X \\ \text{secondo Riemann} \end{matrix}\right) \Longleftrightarrow \left(\begin{matrix} f \text{ continua quasi ovunque} \\ \text{secondo Lebesgue } (2) \end{matrix}\right)$$

Anche le funzioni misurabili si possono caratterizzare in termini di continuità. Premettiamo la seguente definizione.

Sia f una funzione reale definita nel sottoinsime misurabile X di R^m. Si dice che f é quasi continua in X, quando:

$\forall\ \sigma > 0$, esiste un sottoinsieme misurabile X' di X tale che la restrizione di f ad X' sia continua (3) e $m(X - X') < \sigma$.

La definizione é equivalente a quella che si ottiene sostituendo le parole "sottoinsieme misurabile X' " con le parole " sottoinsieme chiuso X' ".

Le funzioni continue quasi ovunque in X secondo Lebesgue sono anche quasi continue, non viceversa (4).

Ciò premesso, sussiste il seguente teorema di Vitali-Lusin:

Per le funzioni f definite in un sottoinsieme X di R^n misurabile (secondo Lebesgue), si ha (5):

(f misurabile) $\Longleftrightarrow$ (f quasi continua) .

---

(1) Ciò vuol dire che sono contigui gli insiemi numerici descritti dalle somme:

$$s_D = \sum_{i=1}^{n} \mu(X_i) \inf f(X_i), \quad S_D = \sum_{i=1}^{n} \mu(X_i) \sup f(X_i) \qquad (\mu = \text{misura di Peano-Jordan})$$

al variare della partizione $D = (X_1, X_2, \ldots, X_n)$ dell'insieme X in insiemi misurabili secondo Peano-Jordan. I predetti insiemi numerici sono in ogni caso separati.

(2) Cioé l'insieme dei punti di discontinuità di f (eventualmente vuoto) ha misura nulla secondo Lebesgue.

(3) Se una funzione f definita in X é continua in $X' \subseteq X$ anche la restrizione di f ad X' é continua, ma non viceversa.

(4) Ad es. la funzione:

$$f: x \in R \longrightarrow \begin{cases} 0 \text{ se } x \text{ é razionale} \\ 1 \text{ se } x \text{ é irrazionale.} \end{cases}$$

é quasi continua, ma non é continua quasi ovunque (in quanto ogni punto di R é di discontinuità per f). E' facile constatare che f é quasi continua: detto X_0 l'insieme dei punti razionali, e posto $X' = R - X_0$, la restrizione di f a X' é continua (in quanto costante) e la misura di $R - X'$, ossia di X_0, é nulla.

(5) Il teorema di Vitali-Lusin si dimostra per le funzioni definite in X misurabile e di misura finita, ma si verifica facilmente che esso sussiste anche se X ha misura infinita. Quest'ultima affermazione sarà dimostrata nell'Appendice [cfr. pag. 329].

Con riferimento ad una funzione reale f definita nel sottoinsieme misurabile X, ivi misurabile, introduciamo ora la seguente definizione.

Si dice che f è essenzialmente limitata (in X), quando esiste un numero positivo c tale che l'insieme (misurabile):

$$\{x \in X : |f(x)| > c\}$$

abbia misura nulla, ossia quando esiste $c > 0$ tale che:

$$|f(x)| \leq c \quad \text{quasi ovunque in } X \quad (1).$$

Esistono funzioni essenzialmente limitate che non sono limitate. Ad es. la funzione:

$$\varphi : x \in [0,1] \longrightarrow \begin{cases} 1 & \text{se } x \neq \frac{1}{n} \quad (n \in N) \\ n & \text{se } x = \frac{1}{n} \end{cases}$$

è essenzialmente limitata, in quanto l'insieme

$$\{x \in [0,1] : |\varphi(x)| > 1\}$$

è costituito dai punti del tipo $1/n$ con $n \in N, n > 1$, e quindi ha misura nulla (in quanto è numerabile). Peraltro φ non è limitata.

Osserviamo che, se X ha misura nulla, ogni funzione reale f definita in X è essenzialmente limitata.

Se f è essenzialmente limitata e $m(X) > 0$, l'insieme:

$$(2) \qquad \{k \in R : f(x) \leq k \ \text{q.o. in } X\}$$

è limitato inferiormente; inoltre tale insieme è dotato di minimo. Il minimo dello insieme (2) si chiama estremo superiore essenziale (o vero estremo superiore) di f (in X), e si denota col simbolo:

$$\text{ess.sup } f .$$

Il numero reale $\ell'' = \text{ess.sup } f$ gode delle seguenti proprietà:

1) $f(x) \leq \ell''$ q.o. in X

2) $\forall \varepsilon > 0$ esiste un sottoinsieme X' di X, misurabile e di misura positiva, tale che $f(x) > \ell'' - \varepsilon \ \forall x \in X'$.

---

(1) Qualunque sia l'insieme misurabile $X \subseteq R^n$, una proprietà $P(x)$ definita in X si dice verificata quasi ovunque in X (abbreviato, q.o. in X) quando l'insieme:

$$\{x \in X : P(x) \text{ è falsa}\}$$

ha misura nulla.

Analogamente , se f é essenzialmente limitata e $m(X) > 0$,
il massimo dell'insieme:

$$\{h \in R : f(x) \geq h \ \text{q.o. in } X\}$$

diciamolo l' , si chiama estremo inferiore essenziale (o vero estremo inferiore) di f , e si denota col simbolo :

$$\text{ess.inf } f .$$

Il numero reale l' gode delle seguenti proprietà :

1) $f(x) \geq l'$ q.o. in X

2) $\forall \varepsilon > 0$ esiste un sottoinsieme X' di X ; misurabile e di misura positiva, tale che $f(x) < l' + \varepsilon \quad \forall x \in X'$.

Nel caso in cui f é limitata, risulta:

$$\sup f \geq \text{ess.sup } f \ , \quad \inf f \leq \text{ess.inf } f .$$

3. L'integrale secondo Lebesgue.

Sia f una funzione misurabile, definita in un sottoinsieme misurabile X di R^n.
In un primo momento, supponiamo f limitata ed X di misura finita.

Per ciascuna partizione finita di X :

$$\Pi = \{X_1, X_2, \dots, X_n\}$$

costituita da insiemi misurabili, consideriamo le somme:

$$(1) \qquad \sum_{i=1}^{n} m(X_i) \inf f(X_i)$$

$$(2) \qquad \sum_{i=1}^{n} m(X_i) \sup f(X_i) \quad .$$

Al variare della partizione Π (e del numero degli elementi di essa) le somme (1) e (2) descrivono due insiemi numerici, che sono separati . In virtù della ipotesi f misurabile , tali insiemi numerici, diciamoli $A(f)$ e $B(f)$, sono altresì contigui: essendo utile vedere come interviene l'ipotesi di misurabilità su f , dimostriamo l'ultima asserzione .

Dim. Essendo f limitata , esiste $\ell > 0$ tale che $|f(x)| \leq \ell \ \forall x \in X$.

Fissato $\varepsilon > 0$, consideriamo una partizione finita dell'intervallo $[-\ell, \ell]$:

$$\{I_1, I_2, \dots, I_k\}$$

costituita da intervalli tutti di ampiezza minore di $\varepsilon / (1 + m(X))$, e poniamo:

$$X_i = f^{-1}(I_i) \qquad \forall i \in \{1, \dots, k\} .$$

Essendo f misurabile, la partizione di X :

$$\{X_1, X_2, \dots, X_k\}$$

é costituita da insiemi misurabili (1) . Poiché:

$$f(X_i) = f\left(f^{-1}(I_i)\right) \subseteq I_i \qquad \forall i \in \{1, \dots, k\} ,$$

si ha , ricordando come sono stati scelti gli intervalli I_i :

---

(1) In quanto X_i é l'intersezione di due insiemi ciscuno dei quali é del tipo (1) del n.2 o del tipo di quelli ivi considerati nella nota a pié di pagina.

$$\sup f(X_i) - \inf f(X_i) \leq m(I_i) < \varepsilon / (1+m(X))$$

e quindi :

$$\sum_{i=1}^{k} m(X_i) \sup f(X_i) - \sum_{i=1}^{k} m(X_i) \inf f(X_i) =$$

$$= \sum_{i=1}^{k} m(X_i) \left[\sup f(X_i) - \inf f(X_i) \right] <$$

$$< \frac{\varepsilon}{1+m(X)} \sum_{i=1}^{k} m(X_i) = \varepsilon \frac{m(X)}{1+m(X)} < \varepsilon .$$

Pertanto gli insiemi $A(f)$ e $B(f)$, che sono separati, sono anche contigui.

L'elemento di separazione degli insiemi $A(f)$ e $B(f)$ si chiama <u>integrale di f esteso ad X secondo Lebesgue</u>, e si denota con uno dei simboli :

(3) $$\int_X f\,dm \quad , \quad \int_X f(x)\,dx .$$

Nell'ipotesi $f \geq 0$, qualunque siano X_1, X_2 parti misurabili di X si ha che:

$$X_1 \subseteq X_2 \Longrightarrow \int_{X_1} f(x)\,dx \leq \int_{X_2} f(x)\,dx .$$

Inoltre, qualunque siano f, g misurabili e limitate in X, risulta:

$$\int_X (c_1 f(x) + c_2 g(x))\,dx = c_1 \int_X f(x)\,dx + c_2 \int_X g(x)\,dx$$

per ogni $c_1, c_2 \in R$.

Eliminiamo ora l'ipotesi che f sia limitata e che X sia di misura finita ferma restando l'ipotesi f misurabile.

Consideriamo un sottoinsieme X' di X, misurabile e di misura finita, nel quale f sia limitata, e consideriamo l'integrale:

$$\int_{X'} |f(x)|\,dx \tag{4}$$

che ha senso secondo la definizione precedentemente introdotta. Al variare di X' tra le parti di X del tipo suddetto, l'integrale (4) descrive un insieme numerico, diciamolo $\mathcal{J}(f)$. Se $\mathcal{J}(f)$ è limitato superiormente, si dice che f <u>è sommabile in X (secondo Lebesgue)</u>. In tale ipotesi:

a) Se $f \geq 0$, si pone:

$$\int_X f(x)\,dx = \sup \mathcal{J}(f).$$

b) Se non è $f \geq 0$, si considerano le funzioni <u>non negative:</u>

$$f^+ = \frac{|f| + f}{2}, \qquad -f^- = \frac{|f| - f}{2}$$

sicché risulta $f^+ + f^- = f$, e $f^+ - f^- = |f|$.

Le funzioni f^+, f^- sono sommabili in X se f, come abbiamo supposto,

è sommabile in X. Si pone allora, per definizione:

$$\int_X f(x)\,dx = \int_X f^+(x)\,dx - \int_X (-f^-)(x)\,dx .$$

Ovviamente se f è limitata e X è di misura finita, la funzione f è sommabile in X, e l'integrale di f introdotto in a) e in b) coincide con quello definito a pag. 39.

Se f e g sono funzioni sommabili in X, anche $f+g$, $f-g$ e λf $\forall \lambda \in R$ sono sommabili in X.

Rileviamo che dall'ipotesi f, g sommabili non consegue che $f\cdot g$ è sommabile. Una condizione sufficiente affinchè $f\cdot g$ sia sommabile è che siano sommabili f^2 e g^2 (1).

Evidentemente ogni funzione reale f, definita in un insieme X_0 di misura nulla, è sommabile in X_0, e risulta:

$$(5) \qquad \int_{X_0} f(x)\,dx = 0 .$$

Osserviamo infine che, se f è definita in un insieme X misurabile secondo Peano-Jordan, ed è sommabile in X secondo Riemann [2], allora f è sommabile anche secondo Lebesgue, e l'integrale di f esteso ad X secondo Riemann è uguale a quello secondo Lebesgue.

4. Funzioni a valori in $\bar{R}$ definite quasi ovunque : sommabilità e integrale secondo Lebesgue.

Sia X un sottoinsieme misurabile di R^n. Denotiamo con Φ la totalità delle funzioni φ ciascuna delle quali è definita in un sottoinsieme X' di X, misurabile e tale che $m(X-X') = 0$, ed è a valori nell'insieme ampliato dei numeri reali.

Qualunque siano $\varphi, \psi \in \Phi$, diremo che φ e ψ sono quasi ovunque uguali, e scriveremo:

$$\varphi = \psi \quad \text{q.o. in } X,$$

quando esiste un sottoinsieme misurabile X_0 di X, di misura nulla, tale

(1) Cioè che f e g siano di quadrato sommabile. L'asserto si deduce dalla disuguaglianza di Schwarz:

$$\int_{X'} |fg|\,dx \le \left(\int_{X'} f^2\,dx \right)^{1/2} \left(\int_{X'} g^2\,dx \right)^{1/2}$$

applicata ad ogni insieme $X' \subseteq X$, misurabile e di misura finita, nel quale $f\cdot g$ sia limitata.

(2) La sommabilità secondo Riemann, e l'integrale secondo Riemann, si definiscono come si è fatto a pag. 40 per la sommabilità e per l'integrale secondo Lebesgue, supponendo X misurabile secondo Peano-Jordan invece che secondo Lebesgue, ed f, invece che misurabile, continua q.o. in X secondo Lebesgue; inoltre gli insiemi X' considerati a pag. 40 si suppongono misurabili secondo Peano-Jordan anzichè secondo Lebesgue. L'integrale (4) avrà senso in virtù del criterio di Vitali-Lebesgue [cfr. (*) pag. 35].

che φ e ψ siano definite in ogni punto di $X - X_0$ e si abbia :

$$\varphi(x) = \psi(x) \quad \forall x \in X - X_0 .$$

La relazione binaria $\mathcal{R}$ in Φ definita come segue:

$$\varphi \mathcal{R} \psi \quad \text{se e solo se} \quad \varphi = \psi \ \text{q.o. in } X$$

è ovviamente una relazione di equivalenza in Φ : ogni elemento dell'insieme quoziente $\Phi / \mathcal{R}$ si chiama una <u>funzione definita quasi ovunque in X</u> . Pertanto una funzione definita quasi ovunque in X è una classe di funzioni appartenenti a Φ , due qualunque delle quali sono quasi ovunque eguali.

Se f è una funzione definita quasi ovunque in X , e gli elementi della classe f verificano quasi ovunque una certa proprietà, si dice che f stessa verifica quasi ovunque quella proprietà. Ad esempio, dire che $f \geq 0$ q.o. in X significa che gli elementi della classe f assumono quasi ovunque valori non negativi . Si dice che f è <u>q.o. finita</u> in X , quando gli elementi della classe f ovunque definiti in X assumono q.o. valori diversi da $\pm \infty$.

Per ogni funzione f definita quasi ovunque in X , ivi q.o. finita, si dice che f è <u>limitata</u> (in X) quando gli elementi della classe f , ovunque definiti in X (*) e a valori reali, sono funzioni essenzialmente limitate in X ;in tale ipotesi , se X ha misura positiva, il vero estremo superiore [risp.inferiore] degli elementi della classe f si chiama <u>estremo superiore</u> [risp. inferiore] di f .

Siano f, g funzioni definite q.o. in X , ivi q.o. finite, e siano φ, ψ elementi rispettivamente della classe f e della classe g , ovunque definiti in X e a valori in R . La classe delle funzioni quasi ovunque eguali a $\varphi + \psi$ è una funzione definita quasi ovunque in X : essa si chiama <u>somma</u> di f e g , e si denota con $f + g$. Analogamente si definiscono $f - g$, $f \cdot g$, $\lambda f \ \forall \lambda \in R$, $|f|^\lambda \ \forall \lambda \in R^+$, nonché f/g se $g \neq 0$ q.o. in X.

Per ogni funzione f definita quasi ovunque in X , ivi q.o. finita, si dice che f è <u>misurabile</u> [risp. <u>sommabile</u>] in X , quando gli elementi della classe f , ovunque definiti in X e a valori in R, sono funzioni misurabili [risp. somma-

(*) Ovviamente esistono sempre rappresentanti di f definiti in tutto X.

bili] in X.

Sia f una funzione definita q.o. in X, sommabile in X. Detto f^* una qualunque degli elementi della classe f, che sia ovunque definito in X ed a valori in R (1), l'integrale di f^* esteso ad X si chiama integrale di f esteso ad X e si denota con il solito simbolo.

Una funzione f definita quasi ovunque in X, ivi q. o. finita, si dice <u>continua</u> (in X) quando esiste un elemento della classe f, ovunque definito in X, che è una funzione continua in X.

5. Proprietà dell'integrale secondo Lebesgue.

Sia f una funzione definita q.o. nel sottoinsieme misurabile X di R^n. Se f è sommabile in X, f è altresì sommabile in ogni parte misurabile Y di X; se inoltre $f \geq 0$ q.o. in X, si ha:

$$0 \leq \int_Y f(x)\,dx \leq \int_X f(x)\,dx .$$

Se f è sommabile in X, risulta:

$$\left| \int_X f(x)\,dx \right| \leq \int_X |f(x)|\,dx$$

$$\int_X \lambda f(x)\,dx = \lambda \int_X f(x)\,dx \quad \forall \lambda \in R .$$

<u>Teorema della media</u>. <u>Se X ha misura finita, e f è misurabile e limitata, risulta:</u>

$$m(X) \inf f(X) \leq \int_X f(x)\,dx \leq m(X) \sup f(X).$$

Se f è sommabile in X, qualunque siano X_1, X_2 parti misurabili di X si ha:

---

(1) Tali elementi della classe f hanno tutti il medesimo integrale, in virtù della (5) del n.3.

$$\int_{X_1 \cup X_2} f(x)dx = \int_{X_1} f(x)dx + \int_{X_2} f(x)dx - \int_{X_1 \cap X_2} f(x)dx ,$$

e in particolare, se X_1 e X_2 sono disgiunti (1):

$$(1) \qquad \int_{X_1 \cup X_2} f(x)dx = \int_{X_1} f(x)dx + \int_{X_2} f(x)dx ;$$

inoltre:

$$\int_{X_1 - X_2} f(x)dx = \int_{X_1} f(x)dx - \int_{X_1 \cap X_2} f(x)dx$$

e quindi, se $X_2 \subseteq X_1$:

$$\int_{X_1 - X_2} f(x)dx = \int_{X_1} f(x)dx - \int_{X_2} f(x)dx .$$

Se f è sommabile in X, per ogni successione $(X_n)_{n \in N}$ di parti di X misurabili e a due a due disgiunte, risulta:

$$(2) \qquad \int_{\bigcup_{n \in N} X_n} f(x)dx = \sum_{n=1}^{\infty} \int_{X_n} f(x)dx .$$

La (1) esprime la <u>finita additività</u>, e la (2) la <u>numerabile additività</u>, dell'integrale di Lebesgue.

Se f e g sono funzioni sommabili in X, qualunque siano $c_1, c_2 \in R$ si ha:

$$\int_X [c_1 f(x) + c_2 g(x)]dx = c_1 \int_X f(x)dx + c_2 \int_X g(x)dx .$$

Nell'ipotesi che X abbia misura finita enunciamo infine la seguente proprietà, detta <u>proprietà di assoluta continuità dell'integrale</u>: per ogni $\varepsilon > 0$ esiste un $\delta > 0$ tale che, per Y sottoinsieme misurabile di X sussiste l'implicazione:

$$m(Y) < \delta \Rightarrow \left| \int_Y f(x)dx \right| < \varepsilon .$$

6. <u>La nozione di misura</u>.

Detto Ω un sottoinsieme misurabile di R^n, denotiamo con $\mathcal{L}_\Omega$ l'insieme costituito dalle parti di Ω misurabili secondo Lebesgue. Una funzione $\mu : \mathcal{L}_\Omega \to [0, +\infty]$ si dice <u>numerabilmente additiva in</u> $\mathcal{L}_\Omega$ quando risulta:

$$\mu\left(\bigsqcup_{n \in N} X_n\right) = \sum_{n=1}^{\infty} \mu(X_n)$$

per ogni successione $(X_n)_{n \in N}$ di elementi di $\mathcal{L}_\Omega$ a due a due disgiunti. Si può dimostrare che una tale funzione μ gode delle stesse proprietà della misura di Lebesgue, e perciò si chiama anch'essa una misura: dunque per <u>misura in</u> $\mathcal{L}_\Omega$ si intende una qualunque funzione definita in $\mathcal{L}_\Omega$, a valori in $[0, +\infty]$ e numerabilmente additiva.

Fissata una funzione $\theta = \theta(x)$ definita q.o. in Ω, non negativa e <u>localmente sommabile in</u> Ω (ossia, sommabile in ogni insieme compatto contenuto in Ω), tenuto conto della numerabile additività dell'integrale di Lebesgue è ovvio che l'applicazione:

$$(1) \qquad \mu : X \in \mathcal{L}_\Omega \longrightarrow \int_X \theta(x)dx \qquad (2)$$

(1) L'uguaglianza sussiste anche se $X_1 \cap X_2 = \emptyset$, intendendo per definizione uguale a zero l'integrale esteso al sottoinsieme vuoto.

(2) Si conviene che l'integrale sia $+\infty$ se θ, misurabile in Ω, non è ivi sommabile.

è una misura in $\mathcal{L}_\Omega$, che si riduce alla misura di Lebesgue nel caso $\Omega=\mathbb{R}^m$ e Θ costante uguale ad 1. La definizione (1) di μ si può formulare mediante l'uguaglianza:

$$\mu(X)=\int_X \Theta(x)\,dx \quad \forall X\in\mathcal{L}_\Omega \tag{2}$$

e la misura si dice dotata di densità, giacché la funzione Θ si chiama la densità di μ.

E' importante evidenziare che esistono misure in $\mathcal{L}_\Omega$ che non sono dotate di densità. Ad esempio, fissato $x_0\in\Omega$, consideriamo la corrispondenza μ che ad ogni $X\in\mathcal{L}_\Omega$ associa 1 se $x_0\in X$ e 0 se $x_0\notin X$: è facile constatare che μ è una misura in $\mathcal{L}_\Omega$, e ovviamente non è dotata di densità, perché se fosse del tipo (2) per $X=\{x_0\}$ si avrebbe $\mu(\{x_0\})=0$, mentre per definizione si ha $\mu(\{x_0\})=1$.

La nozione di misura si estende ad applicazioni che assumono anche valori negativi (misure relative), nonché ad applicazioni definite in insiemi non necessariamente del tipo $\mathcal{L}_\Omega$: ma su ciò non è il caso di soffermarsi in questa sede.

7. Cenni sull'integrale di Stieltjes-Lebesgue.

Tutte le considerazioni svolte nei numeri 3,4 [pag.37-43] si possono ripetere sostituendo la misura m di Lebesgue con un'arbitraria misura μ in $\mathcal{L}_{\mathbb{R}^m}$ (in particolare con la (1) del numero precedente, ponendo ivi $\Omega=\mathbb{R}^m$, ma anche con una misura in $\mathcal{L}_{\mathbb{R}^m}$ non dotata di densità). Per ogni funzione f limitata e misurabile in un insieme X di misura μ finita (ossia tale che $\mu(X)<+\infty$) si viene così ad introdurre l'integrale:

$$\int_X f\,d\mu$$

che si chiama integrale di f rispetto a μ (secondo Stieltjes-Lebesgue); poi si introduce la nozione di funzione sommabile rispetto a μ e il relativo integrale rispetto a μ.

Come si è fatto per l'integrale secondo Lebesgue, le precedenti nozioni si estendono poi a funzioni definite μ-quasi ovunque in un insieme X misurabile secondo Lebesgue.

Se la misura μ è dotata di densità, per la (1) del numero precedente essa gode delle seguenti proprietà:

a) Assume valori finiti sugli insiemi compatti:
$$X\in\mathcal{L}_\Omega \text{ compatto} \Longrightarrow \mu(X)<+\infty.$$

b) E' "assolutamente continua rispetto alla misura di Lebesgue", cioè si ha:
$$m(X)=0 \Longrightarrow \mu(X)=0.$$

Viceversa (teorema di Nikodym) qualunque misura μ avente le proprietà a) e b) è dotata di densità: ossia esiste una funzione Θ definita q.o. in Ω, non negativa e localmente sommabile in Ω, per la quale sussiste la (2) del numero precedente; più in generale, per $X\in\mathcal{L}_\Omega$ ed f misurabile in X sussiste l'equivalenza:

f sommabile in X rispetto a μ $\Leftrightarrow$ f·θ sommabile in X

e risulta:

(1) $$\int_X f\,d\mu = \int_X f(x)\theta(x)\,dx,$$

sicché l'integrale di Stieltjes-Lebesgue [con μ avente le proprietà a) e b)] si può esprimere mediante un integrale di Lebesgue.

In virtù della (1), nelle locuzioni dianzi introdotte si preferisce sostituire le parole "rispetto a μ" con le parole "rispetto alla funzione peso $\theta(x)$"

Poiché il secondo membro della (1) si ottiene formalmente dal primo membro sostituendo in esso $d\mu$ con $\theta(x)dx$, per esprimere sinteticamente che μ è la corrispondenza (1) di pag. 44 si scrive l'uguaglianza $d\mu = \theta(x)dx$

8. Funzioni a variazione limitata. Funzioni assolutamente continue.

Sia $\alpha(x)$ una funzione reale della variabile reale x, definita in un intervallo $[a,b]$.

Si dice che α è a variazione limitata quando esiste una costante positiva K tale che, qualunque sia $n \in N$ e qualunque siano i punti:

$$x_0 = a < x_1 < x_2 < \dots < x_n = b$$

si abbia:

$$\sum_{i=1}^{n} |\alpha(x_i) - \alpha(x_{i-1})| \leq K .$$

Sussistono i teoremi:

1. Una funzione è a variazione limitata se e solo se è la differenza di due funzioni crescenti.
2. Una combinazione lineare di funzioni a variazione limitata è a sua volta una funzione a variazione limitata.
3. Ogni funzione a variazione limitata è derivabile quasi ovunque, e la sua derivata è una funzione sommabile.

Ovviamente una funzione α crescente è a variazione limitata (ed è la differenza delle funzioni crescenti α e 0). Osserviamo che dal teorema 1 si deduce l'equivalenza:

α è a variazione limitata $\Leftrightarrow$ α è la differenza di due funzioni monotone

DIM. L'implicazione $\Rightarrow$ segue dal teor. 1. Per dimostrare $\Leftarrow$ supponiamo $\alpha = \alpha_1 - \alpha_2$, con α_1, α_2 monotone. Allora α è a variazione limitata: se α_1, α_2 sono entrambe crescenti, per il teor. 1; se sono entrambe decrescenti, ancora per il teor. 1 dato che $\alpha = (-\alpha_2) - (-\alpha_1)$; se α_1 è crescente e α_2 decrescente, nuovamente per il teor. 1 dato che $\alpha = [\alpha_1 + (-\alpha_2)] - 0$ e la funzione $[\alpha_1 + (-\alpha_2)]$ è crescente; infine, se α_1 è decrescente e α_2 crescente, perché α è l'opposta di una funzione a variazione limitata.

Una funzione reale $\alpha(x)$ definita in $[a,b]$ si dice assolutamente continua (brevemente ass. cont.) quando è verificata la seguente condizione:

Per ogni $\varepsilon > 0$ esiste un $\delta_\varepsilon > 0$ tale che, qualunque siano $n \in \mathbb{N}$ e $[a_1, b_1], \dots, [a_n, b_n]$ intervalli contenuti in $[a,b]$, a due a due disgiunti per $n \neq 1$, si abbia:

$$\sum_{i=1}^{n} (b_i - a_i) < \delta_\varepsilon \Rightarrow \sum_{i=1}^{n} |f(b_i) - f(a_i)| < \varepsilon .$$

Evidentemente ogni funzione lipschitziana è assolutamente continua.

Sussistono i teoremi:

4. Ogni funzione assolutamente continua è a variazione limitata.

5. Se f è una funzione sommabile in (a,b), la funzione

$$\alpha(x) = \int_a^x f(t)\,dt \tag{1}$$

è assolutamente continua in $[a,b]$, e risulta $\alpha' = f$ q.o. in (a,b).

Sia f una funzione reale definita q.o. in (a,b). Si chiama primitiva generalizzata di f (brevemente pr. gen. di f) ogni funzione F continua in $[a,b]$, derivabile q.o. in (a,b) e tale che $F' = f$.

Sia ora f sommabile in (a,b). La (1) è una pr. gen. di f, ed è ass. cont., ma non tutte le pr. gen. di f sono ass. cont.: infatti si può costruire [5) pag.(11)] una funzione φ, continua in $[a,b]$ ma non ass. cont., derivabile q.o. in (a,b) e tale che $\varphi' = 0$ q.o. in (a,b); ovviamente la funzione $\beta = \alpha + \varphi$ è una pr. gen. di f non ass. cont. La φ non è costante (altrimenti sarebbe ass. cont.), sicché β ed α sono due pr. gen. di f la cui differenza (ossia φ) non è costante.

Invece si dimostra che è costante la differenza di due pr. gen. di f che siano ass. cont. Ne segue che qualunque sia G pr. gen. di f, purché assolutamente continua, risulta $\int_a^b f(x)\,dx = G(b) - G(a)$.

9. Integrale di Stieltjes di una funzione continua di una variabile.

Fissata una funzione reale α definita in un intervallo $[a,b]$, ivi a variazione limitata, per ogni funzione reale f continua in $[a,b]$ si definisce "l'integrale di Stieltjes", che si denota col simbolo:

(1) $$\int_a^b f(x)\,d\alpha$$

o più concisamente con

(2) $$\int_a^b f\,d\alpha \ .$$

Per il teor. 3 di pag. 46 ha significato l'integrale:

(3) $$\int_a^b f(x)\alpha'(x)\,dx$$

ma questo, pur essendo $d\alpha = \alpha'(x)dx$ q.o. in $[a,b]$, in generale non coincide con l'integrale (1). Vedremo che gli integrali (1) e (3) sono uguali qualunque sia f solo quando la funzione α è assolutamente continua in $[a,b]$.

Per definire l'integrale di Stieltjes supponiamo in un primo momento che α, non necessariamente continua, sia crescente in $[a,b]$.

Per ogni $n \in \mathbb{N}-\{1\}$ decomponiamo $[a,b]$ in n intervalli, mediante i punti:

$$x_0 = a < x_1 < x_2 < \dots < x_{n-1} < x_n = b$$

scelti arbitrariamente, e consideriamo le somme:

(4) $$\sum_{i=0}^{n-1} m_i\,[\alpha(x_{i+1}) - \alpha(x_i)] \quad , \quad \sum_{i=0}^{n-1} M_i\,[\alpha(x_{i+1}) - \alpha(x_i)]$$

$[x_i, x_{i+1}]$. Il lettore osservi che nel caso particolare $\alpha(x) = x$ le (4) sono le somme di Riemann relative alla funzione continua f : orbene, come accade nel caso particolare $\alpha(x) = x$, le somme (4) descrivono due insiemi numerici separati e contigui (omettiamo la dimostrazione). L'elemento di separazione dei due insiemi, che nel caso particolare anzidetto è l'integrale di f esteso ad $[a,b]$, si chiama integrale di Stieltjes di f rispetto ad α, esteso ad $[a,b]$, e si denota con il simbolo (3).

La nozione appena introdotta si estende subito al caso che la funzione continua f sia a valori complessi. Se $f = f_1 + j f_2$ si pone per definizione:

$$\int_a^b f\,d\alpha = \int_a^b f_1\,d\alpha + j \int_a^b f_2\,d\alpha \ .$$

Passiamo ora al caso generale che α sia una funzione a variazione limitata. Siccome una tale funzione è la differenza di due funzioni crescenti [teor. 1 pag. 46], se $\alpha = \alpha_1 - \alpha_2$ con α_1, α_2 crescenti si pone per definizione:

$$\int_a^b f\,d\alpha = \int_a^b f\,d\alpha_1 - \int_a^b f\,d\alpha_2 \ .$$

Osserviamo infine che l'integrale di Stieltjes si definisce anche nel caso che la funzione a variazione limitata α sia a valori complessi. Una funzione

complessa α definita in $[a,b]$ si dice a variazione limitata quando tali sono la parte reale φ e il coefficiente ψ della parte immaginaria. Se $f=\varphi+j\psi$ con φ,ψ a variazione limitata, si pone per definizione:

$$\int_a^b f\,d\alpha = \int_a^b \varphi\,d\alpha + j\int_a^b \psi\,d\alpha \ .$$

Nel seguito occorrerà utilizzare la proprietà additiva dell'integrale di Stieltjes:

$$(5) \qquad \int_a^b f\,d\alpha = \int_a^c f\,d\alpha + \int_c^b f\,d\alpha \qquad \forall c\in\,]a,b[\ .$$

Non è escluso che la funzione α, definita in $[a,b]$, sia costante nell'intervallo (a,b) chiuso, aperto o semiaperto. Osserviamo che:

1. Qualunque sia la funzione f reale o complessa, continua in $[a,b]$, e qualunque sia α reale o complessa, definita in $[a,b]$, si ha:

$$(6) \qquad \alpha \text{ costante in } [a,b] \Rightarrow \int_a^b f\,d\alpha = 0$$

$$(7) \qquad \alpha \text{ costante in } [a,b[\Rightarrow \int_a^b f\,d\alpha = f(b)[\alpha(b)-\alpha(a)]$$

$$(8) \qquad \alpha \text{ costante in }]a,b] \Rightarrow \int_a^b f\,d\alpha = f(a)[\alpha(b)-\alpha(a)]$$

Il caso α costante in $]a,b[$ si riconduce ai casi (7), (8) mediante la (5).

DIM. Ovviamente basta dimostrare l'asserto nel caso che f ed α siano reali, o che α sia crescente (il termine "crescente" fin qui adoperato non è inteso in senso stretto): in tali ipotesi l'integrale (1) è l'elemento di separazione delle somme (4).

Se α è costante in $[a,b]$, gli addendi delle somme (4) sono tutti nulli e l'asserto è banale.

Se α è costante in $[a,b[$, solo l'ultimo addendo delle due somme non è necessariamente nullo: quello della prima somma è $m_{n,n-1}[\alpha(b)-\alpha(a)]$, quello della seconda somma è $M_{n,n-1}[\alpha(b)-\alpha(a)]$. Siccome $m_{n,n-1}\le f(x)\le M_{n,n-1}\ \forall x\in[a,b]$, stante la continuità di f in b risulta $\lim_{n\to\infty} m_{n,n-1}=f(b)$ e $\lim_{n\to\infty} M_{n,n-1}=f(b)$: dopo ciò è chiaro che l'elemento di separazione delle due somme è $f(b)[\alpha(b)-\alpha(a)]$.

Se α è costante in $]a,b]$, il ragionamento è del tutto analogo.

Osserviamo che se α è costante in $[a,b[$ o in $]a,b]$, ma non nell'intervallo chiuso $[a,b]$, la funzione derivata α' è nulla quasi ovunque in (a,b) e perciò l'integrale (3) è uguale a zero. Invece l'integrale (1), per il teorema appena stabilito, non è necessariamente nullo [nel caso (7) è $\neq 0$ se $f(b)\neq 0$, nel caso (8) è $\neq 0$ se $f(a)\neq 0$]. Notiamo che la funzione α, banalmente a variazione limitata, non è assolutamente continua in $[a,b]$ dato che non è continua

in b nel caso (7), in a nel caso (8). Peraltro si dimostra che:

2. Qualunque sia la funzione f reale o complessa, continua in $[a,b]$, e qualunque sia la funzione α reale o complessa, definita in $[a,b]$, se α è assolutamente continua in $[a,b]$ risulta:

$$(9) \qquad \int_a^b f\,d\alpha = \int_a^b f(x)\alpha'(x)dx$$

Allo scopo di chiarire il motivo per cui, come osservato con un esempio prima del Teor.2, la (9) non sussiste quando α è a variazione limitata ma non assolutamente continua, supponiamo che sia $\alpha' \geq 0$ q.o. in $[a,b]$. In tale ipotesi la funzione μ, definita nell'insieme $\mathcal{L}_{[a,b]}$ delle parti di $[a,b]$ misurabili secondo Lebesgue ponendo:

$$\mu(X) = \int_X \alpha'(x)dx \qquad \forall X \in \mathcal{L}_{[a,b]}$$

è una misura in $\mathcal{L}_{[a,b]}$; se α è assolutamente continua in $[a,b]$, qualunque siano $x_1, x_2 \in [a,b]$ con $x_1 < x_2$ risulta:

$$\mu([x_1,x_2]) = \alpha(x_2) - \alpha(x_1)$$

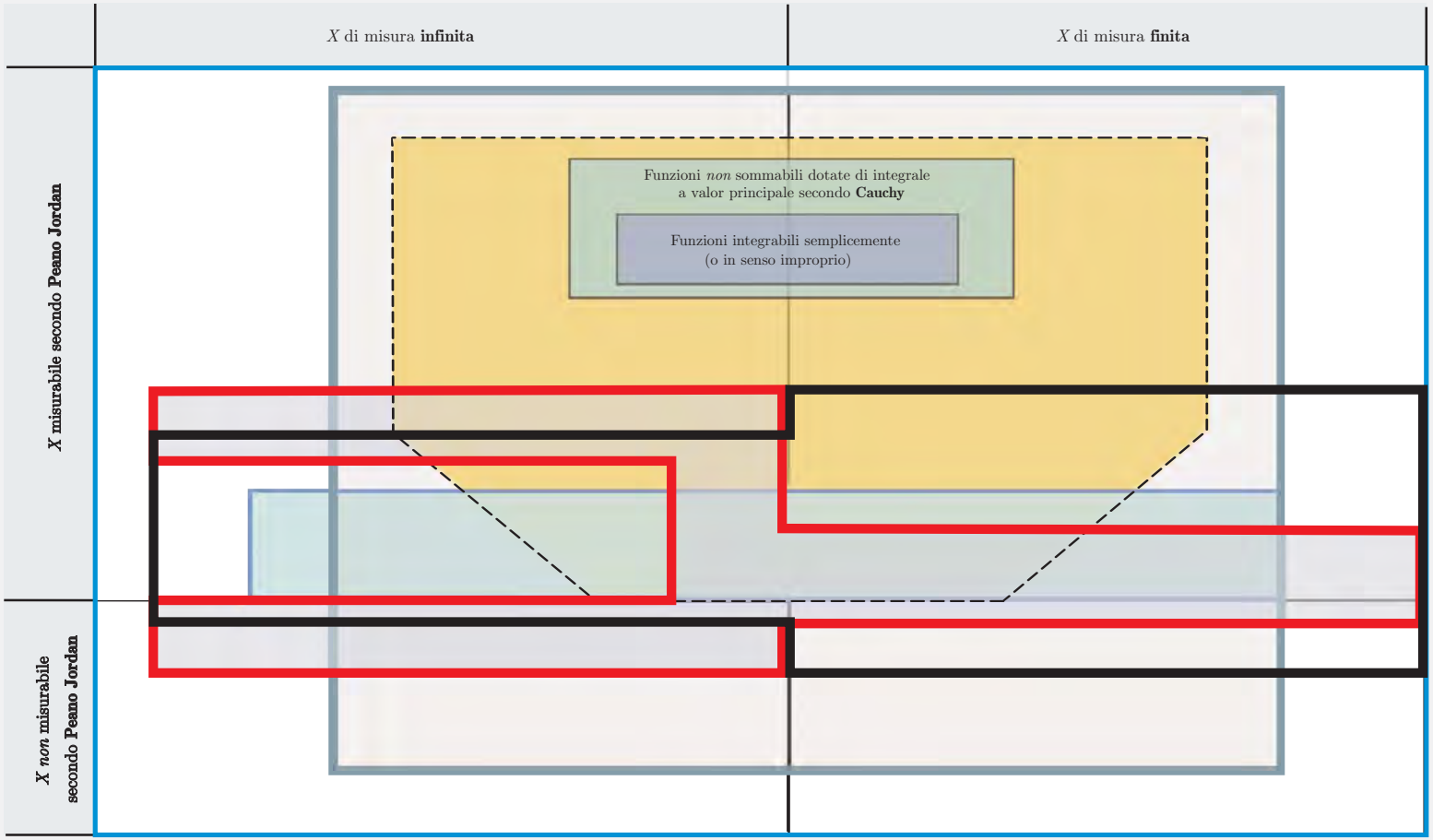
sicché la restrizione di μ all'insieme degli intervalli chiusi contenuti in $[a,b]$ coincide con la cosiddetta funzione di intervallo generata da α. È questa coincidenza che comporta la (9) (su ciò non ci soffermiamo).

Se α non è assolutamente continua, si ha

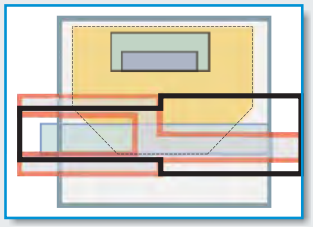
$$\mu([x_1,x_2]) = \int_{[x_1,x_2]} \alpha'(x)dx$$

ma il secondo membro non è uguale ad $\alpha(x_2) - \alpha(x_1)$: la funzione α, se continua, è una primitiva generalizzata di α' ma non sussiste la formula fondamentale del calcolo integrale, in quanto la differenza di due primitive generalizzate non è necessariamente costante [pag. 47 dopo il teor. 5].

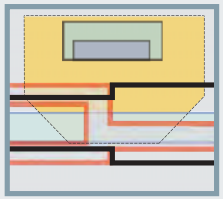
SCHEMA RIASSUNTIVO SULLA RIPARTIZIONE DELLA CLASSE DELLE FUNZIONI MISURABILI IN $X \subseteq \mathbb{R}^n$



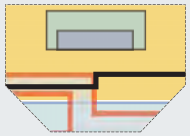
LEGENDA



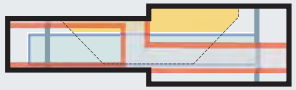
Funzioni definite q.o. in $X \subseteq \mathbb{R}^n$ misurabile secondo **Lebesgue**, ivi misurabili ossia quasi continue.



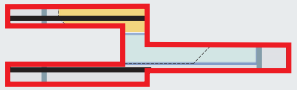
Funzioni continue quasi ovunque in X secondo **Lebesgue**.



Funzioni generalmente continue in un intervallo X di $\mathbb{R}$.



Funzioni sommabili



Funzioni misurabili limitate

Funzioni sommabili (o assolutamente integrabili) secondo **Riemann**. Per X di misura finita l'integrale si può definire come limite delle somme di **Cauchy-Riemann**, cioè con il procedimento che abitualmente si segue nel caso di f continua in X compatto.

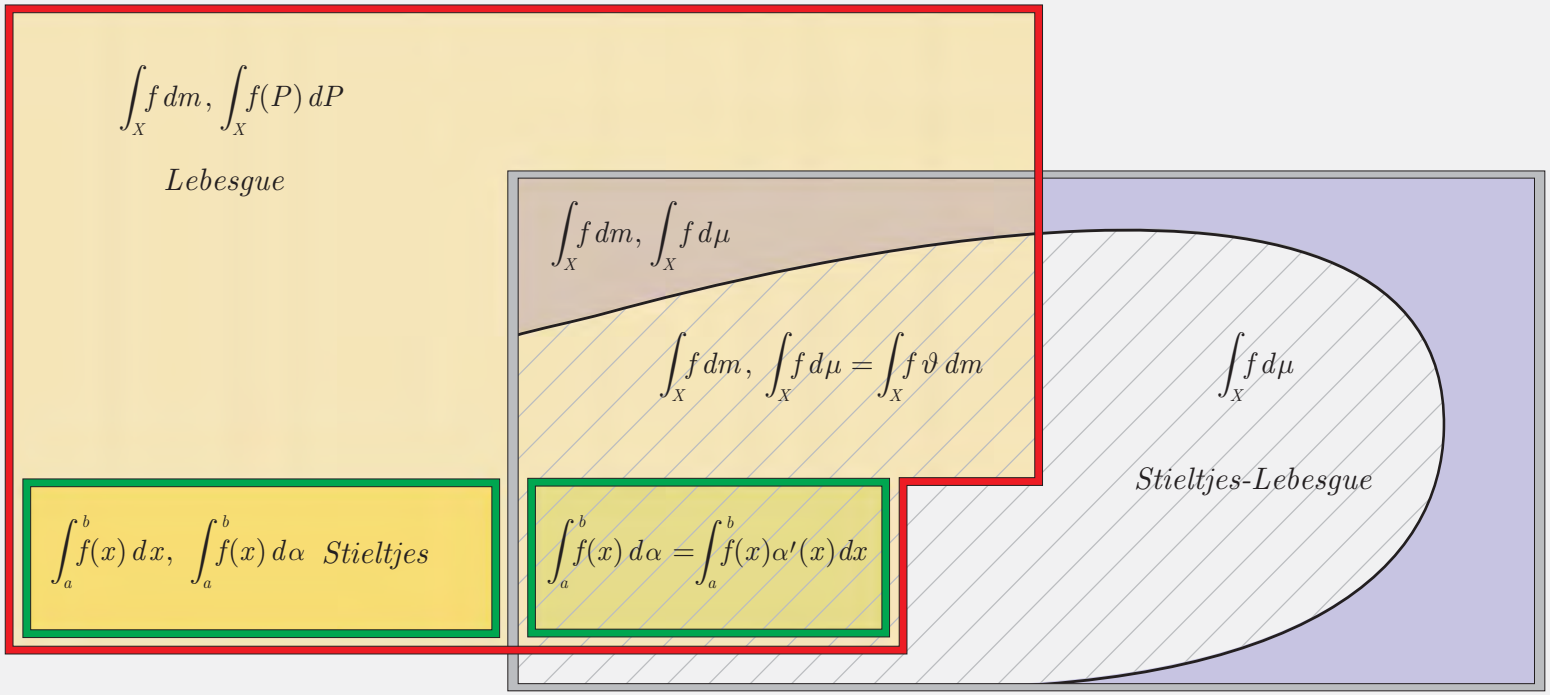
SCHEMA RIASSUNTIVO SULLA NOZIONE DI INTEGRALE

Notazioni.

$f =$ funzione misurabile definita q.o. in $X \subseteq \mathbb{R}^n$ $(n \geqslant 1)$ misurabile secondo Lebesgue;

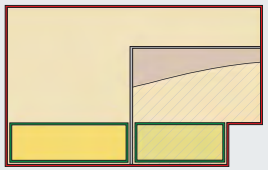
$m =$ misura di Lebesgue; $\mu =$ misura in $\mathcal{L}_X$ (insieme delle parti di X misurabili secondo Lebesgue).

$\alpha =$ funzione a variazione limitata $(n = 1,\ X = [a,b])$

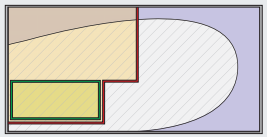


N.B. L'integrale $\int_X f\,d\mu$ si definisce anche in uno spazio mensurale $(S, \mathcal{M}, \mu)$: $S =$ insieme, $\mathcal{M} = \sigma$–algebra di parti di S, $\mu =$ misura definita in $\mathcal{M}$. Qui $S = X$, $\mathcal{M} = \mathcal{L}_X$.

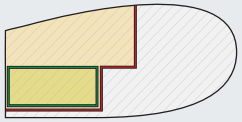
LEGENDA



f sommabile rispetto ad m



f sommabile rispetto a μ



f sommabile rispetto a μ, μ dotata di densità

caso $n = 1$

$X = [a,b]$, f continua in X, α a variazione limitata.

α non assolutamente continua; l'integrale di Stieltjes non è un integrale di Stieltjes-Lebsgue.

α assolutamente continua quindi $d\alpha = \alpha'(x)dx = \alpha' dm = d\mu$

Capitolo III

Alcuni richiami di Algebra. Operatori e funzionali lineari.

1. Spazi vettoriali.

Con la lettera K denotiamo indifferentemente il campo R dei numeri reali o il campo C dei numeri complessi.

Ciò premesso, siano: S un insieme, τ una legge di composizione interna in S, ossia un'applicazione di $S \times S$ in S, e θ una legge di composizione esterna tra K ed S, ossia un'applicazione di $K \times S$ in S.

Per ogni $(x,y) \in S \times S$, l'elemento $\tau(x,y)$ di S si denota con $x \tau y$: per semplicità, scriveremo $x+y$ in luogo di $x \tau y$; analogamente, per ogni $(\lambda, x) \in K \times S$, l'elemento $\theta(\lambda, x)$ di S sarà denotato con λx.

Supporremo che la coppia (S,τ) sia un gruppo abeliano: si dice che (S,τ) è un gruppo quando sono verificate le seguenti condizioni:

a) τ è associativa, ossia risulta:

$$x+(y+z) = (x+y)+z \qquad \forall x,y,z \in S.$$

b) τ è dotata di elemento neutro: ossia, esiste un(unico) elemento di S, che denotiamo con 0, tale che:

$$x+0 = 0+x = x \qquad \forall x \in S.$$

c) ogni elemento x di S è dotato di simmetrico rispetto a τ: ossia, per ogni $x \in S$ esiste un(unico) elemento x' di S tale che:

$$x+x' = x'+x = 0.$$

Il gruppo (S,τ) si dice abeliano quando la legge τ è commutativa, ossia quando risulta:

$$x+y = y+x \qquad \forall x,y \in S.$$

Nell'ipotesi che la coppia (S,τ) sia un gruppo abeliano, si dice che la terna (S,τ,θ) è uno spazio vettoriale sul campo K (di sostegno S) quando sono verificate le seguenti condizioni, qualunque siano $\lambda, \mu \in K$ e $x,y \in S$:

$$\lambda(x+y) = \lambda x + \lambda y \quad , \quad (\lambda+\mu)x = \lambda x + \mu x$$
$$\lambda(\mu x) = (\lambda\mu)x \quad , \quad 1x = x \quad .$$

Per semplicità, con la lettera S indicheremo, oltre che l'insieme S, anche la terna (S, τ, θ); inoltre, poichè considereremo sempre spazi vettoriali sul campo K, li chiameremo semplicemente "spazi vettoriali". Quando occorrerà precisare che $K = R$ [che $K = C$], diremo "spazio vettoriale sul campo reale" o più brevemente "spazio vettoriale reale" ["spazio vettoriale sul campo complesso", o "spazio vettoriale complesso"]. Se τ e θ sono rispettivamente l'operazione di addizione e l'operazione di moltiplicazione in R [risp. in C], lo spazio vettoriale (R, τ, θ) [risp. (C, τ, θ)] sarà denotato con K.

Qualunque sia lo spazio vettoriale S, gli elementi di S si chiamano <u>vettori</u>, e gli elementi del campo K <u>scalari.</u>

Consideriamo ora uno spazio vettoriale S, e introduciamo le seguenti definizioni. Per ogni $n \in N$, qualunque siano i vettori $x_1, \dots, x_n$ di S e gli scalari $\lambda_1, \lambda_2, \dots, \lambda_n$, il vettore:

$$\lambda_1 x_1 + \lambda_2 x_2 + \dots + \lambda_n x_n$$

si chiama <u>combinazione lineare dei vettori</u> $x_1, x_2, \dots, x_n$ <u>di coefficienti</u> $\lambda_1, \dots, \lambda_n$. I vettori $x_1, \dots, x_n$ si <u>dicono linearmente indipendenti</u> quando la relazione

$$\lambda_1 x_1 + \lambda_2 x_2 + \dots + \lambda_n x_n = 0$$

è verificata (se e) soltanto se $\lambda_1 = \lambda_2 = \dots = \lambda_n = 0$. Se n vettori non sono linearmente indipendenti, essi si dicono <u>linearmente dipendenti</u>.

Si dice che i <u>vettori della famiglia</u> $(x_i)_{i \in I}$ sono <u>linearmente indipendenti</u> quando per ogni sottoinsieme finito $\{i_1, \dots, i_n\}$ di I, i vettori $x_{i_1}, \dots, x_{i_n}$ sono linearmente indipendenti.

Qualunque sia $L \subseteq S$, si dice che L è una <u>parte libera</u> di S quando ogni sottoinsieme finito di L è costituito da vettori linearmente indipendenti. Una parte libera L di S si dice una <u>base</u> dello spazio vettoriale S, quando ciascun elemento di S è una combinazione lineare (di un numero finito) di elementi di L. Si dimostra che <u>ogni spazio vettoriale costituito da più di un elemento è dotato di base</u>.

Se esiste una base finita di S, si dice che S è uno spazio <u>a dimensione finita</u>: in tale ipotesi, tutte le basi di S sono finite ed hanno lo stesso numero di elementi, che si chiama <u>la dimensione</u> di S. Se non esiste alcuna base finita, si dice che lo spazio <u>è a dimensione infinita</u>, oppure che <u>ha dimensione infinita</u>.

Si conviene che gli spazi vettoriali costituiti da un solo elemento abbiano dimensione 0.

Sia H una parte non vuota di S avente la seguente proprietà: ogni combinazione lineare di elementi di H appartiene ad H (1). Ciò implica che la restrizione ad $H \times H$ della legge di composizione interna τ in S, diciamola τ_H, é una legge di composizione interna in H; inoltre, la restrizione a $K \times H$ della legge di composizione esterna θ tra K ed S, diciamola θ_H, é una legge di composizione esterna tra K ed H. E' immediato constatare che la terna (H, τ_H, θ_H) é uno spazio vettoriale, che si chiama sottospazio vettoriale di S di sostegno H. Brevemente, diremo che H é un sottospazio di S, purché non vi sia possibilità di equivoci. Se S ha dimensione finita n, ogni sottospazio di S ha dimensione finita $m \leq n$.

Assegnato un sottoinsieme S_0 di S, sia H l'insieme costituito dagli elementi di S ciascuno dei quali é una combinazione lineare di elementi di S_0: evidentemente H é un sottospazio di S, che dicesi generato da S_0; si dice anche che S_0 é un sistema di generatori di H. Naturalmente, se S_0 é una parte libera di S, e se il sottospazio generato da S_0 coincide con S, S_0 é una base di S.

Qualunque siano E_1, E_2 parti di S, col simbolo:

$$E_1 + E_2$$

si denota l'insieme costituito dagli elementi di S ciascuno dei quali é del tipo $x_1 + x_2$ con $x_1 \in E_1$ e $x_2 \in E_2$; se E_1 é costituito da un unico elemento, diciamolo x_1, l'insieme $\{x_1\} + E_2$ si denota con $x_1 + E_2$.

Qualunque siano $E_0 \subseteq S$ e $K_0 \subseteq K$, col simbolo:

$$K_0 E_0$$

si denota l'insieme costituito dagli elementi di S ciascuno dei quali é del tipo λx con $\lambda \in K_0$ e $x \in E_0$; se K_0 [risp. E_0] é costituito da un unico elemento, diciamolo λ_0 [risp. x_0], in luogo di $\{\lambda_0\} E_0$ [risp. in luogo di $K_0 \{x_0\}$] si scrive $\lambda_0 E_0$ [risp. $K_0 x_0$].

---

(1) Equivalentemente, supponiamo che sussistano le seguenti implicazioni:
$(x \in H \text{ e } y \in H) \Rightarrow (x+y \in H)$; $(\lambda \in K \text{ e } x \in H) \Rightarrow (\lambda x \in H)$.

Siano H_1, H_2 sottospazi di S. Si dice che S é somma diretta di H_1 e H_2, quando risulta:

$$H_1 + H_2 = S \quad , \quad H_1 \cap H_2 = \{0\} \quad ,$$

cioé quando per ogni $x \in S$ esistono, e sono univocamente determinati, $h_1 \in H_1$ e $h_2 \in H_2$ tali che $x = h_1 + h_2$.

Se S é somma diretta di H_1 e H_2, i sottospazi H_1 e H_2 si dicono supplementari; si dice anche che ciascuno di essi e' il supplementare dell'altro.

Per ogni sottospazio H di S, e per ogni $x_0 \in S$, il sottoinsieme $x_0 + H$ dicesi una varietà lineare di S. Qualunque siano $x_0, x_1 \in S$, con $x_1 \neq 0$, la varietà lineare di S:

$$x_0 + K x_1$$

si chiama una retta di S (passante per x_0); in particolare, la retta Kx_1 e' un sottospazio di S. Detto I l'intervallo $[0,1]$, il sottoinsieme della retta $x_0 + K(x_1 - x_0)$:

$$x_0 + I(x_1 - x_0)$$

si chiama segmento di S, di estremi x_0 ed x_1; pertanto il segmento di S di estremi x_0 ed x_1 e' l'insieme costituito dai punti di S del tipo:

$$(1-t)x_0 + t x_1 \quad \text{con } t \in [0,1]$$

Un sottospazio H di S si chiama un iperpiano di S quando e' il supplementare di una retta passante per 0: ossia, quando esiste $x_1 \neq 0$ tale che i sottospazi H e Kx_1 siano supplementari. Se il sottospazio H e' un iperpiano di S, per ogni $x_0 \in S - H$ anche la varietà lineare $x_0 + H$ si chiama un iperpiano di S.

Un sottoinsieme E di S si dice convesso quando, qualunque siano $x', x'' \in E$, il segmento di estremi x', x'' é contenuto in E.
Evidentemente, qualunque siano E_1, E_2, E parti di S, si ha:

$$E_1, E_2 \text{ convessi} \implies E_1 + E_2 \text{ convesso}$$

$$E \text{ convesso} \implies \lambda E \text{ convesso } \forall \lambda \in K$$

Evidente é altresì che l'intersezione, supposta non vuota, di una qualunque famiglia di parti convesse di S é una parte convessa di S: pertanto, per ogni sottoinsieme S_0 di S esiste il più piccolo (rispetto alla relazione di inclusione) insieme convesso contenente S_0: tale insieme, che si chiama l'inviluppo convesso di S_0, é costituito dagli elementi di S del tipo:

$$\lambda_1 x_1 + \lambda_2 x_2 + \cdots + \lambda_n x_n$$

con $n \in N$, $x_1, \ldots, x_n \in S_0$, $\lambda_1, \ldots, \lambda_n$ numeri reali tali che $\lambda_i > 0$, $\sum_{i=1}^{n} \lambda_i = 1$.

Sia H un sottospazio di S. La relazione binaria $\mathcal{R}$ in S definita come segue:

$$x \mathcal{R} y \text{ se e solo se } x - y \in H$$

é una relazione di equivalenza in S : per ogni $x \in S$, la classe di equivalenza di x rispetto ad $\mathcal{R}$ si chiama classe di equivalenza di x modulo H; l'insieme quoziente $S/\mathcal{R}$ si chiama insieme quoziente di S rispetto ad H, e si denota con S/H .

Siano X, Y elementi di S/H . Comunque si scelgano $x \in X$ e $y \in Y$, la classe di equivalenza modulo H di $x+y$ é sempre la stessa : denotiamola con $X+Y$; analogamente, per ogni $\lambda \in K$ e per ogni $X \in S/H$, comunque si scelga $x \in X$ la classe di equivalenza di λx é sempre la stessa: denotiamola con λX . In tal modo risultano definite una legge di composizione interna τ_H in S/H , e una legge di composizione esterna θ_H tra K ed S/H. Si verifica facilmente che la terna $(S/H, \tau_H, \theta_H)$ é uno spazio vettoriale, che si chiama spazio vettoriale quoziente di S rispetto ad H : per semplicità, tale spazio vettoriale si denota semplicemente con S/H .

Sia $(S_1, S_2, \ldots, S_n)$ una n-pla di spazi vettoriali, e sia $S = S_1 \times \cdots \times S_n$. Ponendo, qualunque siano $x = (x_1, \ldots, x_n) \in S$ e $y = (y_1, \ldots, y_n) \in S$:

$$x + y = (x_1 + y_1, \ldots, x_n + y_n)$$

si definisce una legge di composizione interna τ in S ; ponendo inoltre, qualunque siano $\lambda \in K$ e $x = (x_1, \ldots, x_n) \in S$:

$$\lambda x = (\lambda x_1, \ldots, \lambda x_n) ,$$

si definisce una legge di composizione esterna θ tra K ed S . Si verifica facilmente che la terna (S, τ, θ) é uno spazio vettoriale, che si chiama spazio vettoriale prodotto della n-pla di spazi vettoriali $(S_1, \ldots, S_n)$, e si denota semplicemente con $S_1 \times S_2 \times \cdots \times S_n$.

Se $S_1 = S_2 = \cdots = S_n$, lo spazio vettoriale prodotto si denota con S^n. In particolare per $S = K$ si ottiene lo spazio vettoriale K^n.

ESEMPI DI SPAZI VETTORIALI

1. Lo spazio $L^p(\Omega)$ [p reale ≥ 1 [1]; Ω aperto di R^k].

E' costituito dalle funzioni reali definite q.o. in Ω:

$$x = x(t) \quad , \text{ con } t = (t_1, \dots t_k) \, ,$$

tali che $|x(t)|^p$ sia sommabile in Ω [per $p=2$, di quadrato sommabile in Ω].

Le operazioni sono quelle usuali: se x_0 e y_0 sono rappresentanti definiti in tutto Ω rispettivamente delle classi x ed y appartenenti ad $L^p(\Omega)$, la somma $x+y$ e' la classe rappresentata dalla funzione x_0+y_0 [cioe' dalla funzione che ad ogni $t \in \Omega$ associa $x_0(t)+y_0(t)$], e λx e' per ogni $\lambda \in R$ la classe rappresentata dalla funzione λx_0 [cioe' dalla funzione che ad ogni $t \in \Omega$ associa $\lambda x_0(t)$]. E' ovvio che per $\lambda \in R$ e $x \in L^p(\Omega)$ si ha $\lambda x \in L^p(\Omega)$; il fatto che $x \in L^p(\Omega), y \in L^p(\Omega) \Rightarrow x+y \in L^p(\Omega)$ si deduce dalla disuguaglianza:

$$(1) \qquad (a+b)^p \leq 2^p (a^p + b^p) \quad \forall a, b \in R_+^+ \quad [2].$$

Se p e p' sono esponenti coniugati, cioe' tali che:

$$\frac{1}{p} + \frac{1}{p'} = 1 \, ,$$

sussiste l'implicazione:

$$(2) \qquad x \in L^p(\Omega) \, , \, y \in L^{p'}(\Omega) \Rightarrow xy \in L^1(\Omega) \, ,$$

che si deduce in modo ovvio dalla disuguaglianza di Schwarz-Holder:

$$(3) \qquad \int_\Omega |x(t) y(t)| \, dt \leq \left[\int_\Omega |x(t)|^p dt \right]^{\frac{1}{p}} \left[\int_\Omega |y(t)|^{p'} \right]^{\frac{1}{p'}}$$

valida qualunque siano $x(t), y(t)$ funzioni misurabili in Ω [3]. Tale disuguaglianza e' una generalizzazione della disuguaglianza di Schwarz, che si ritrova per $p=p'=2$.

Se Ω ha misura finita, la famiglia di spazi $(L^p(\Omega))_{p \in [1, +\infty[}$ e' decrescente rispetto alla relazione di inclusione. Si ha cioe': $p_1 < p_2 \Rightarrow L^{p_1}(\Omega) \supset L^{p_2}(\Omega)$.

Quanto sopra si può ripetere considerando funzioni complesse anziche' funzioni reali. Si ottiene cosi' lo spazio $L^p(\Omega)$ complesso.

2. Lo spazio $L^\infty(\Omega)$ [Ω aperto di R^k].

E' costituito dalle funzioni reali definite q.o. in Ω, ivi misurabili e limitate. Le operazioni sono quelle usuali, gia' richiamate nel precedente esempio 1. Se Ω ha misura finita, $L^\infty(\Omega)$ e' un sottospazio di $L^p(\Omega)$ qualunque sia $p \geq 1$, ed e' contenuto strettamente nell'intersezione di tutti gli $L^p(\Omega)$ [ad es. la funzione $x = \log t$ appartiene ad $L^p(]0,1[)$ $\forall p \geq 1$, ma non e' limitata e quindi non appartiene ad $L^\infty(]0,1[)$].

---

[1] L'ipotesi occorre per la topologia che sara' associata allo spazio. Il motivo sara' chiaro nel seguito.

[2] Se $a \leq b$ si ha $(a+b)^p \leq (b+b)^p = 2^p b^p \leq 2^p (b^p + a^p)$; se $a > b$ si ha $(a+b)^p < (a+a)^p = 2^p a^p \leq 2^p (a^p + b^p)$.

[3] Di una funzione misurabile in Ω, ivi non negativa, e' consuetudine considerare l'integrale senza supporre che la funzione sia sommabile, con la convenzione che l'integrale sia $+\infty$ se la funzione non e' sommabile.

Se si considerano funzioni complesse invece che funzioni reali, si ottiene lo spazio $L^\infty(\Omega)$ complesso.

3. Lo spazio $L^p(\Omega,\mu)$ [$p\in[1,+\infty]$, Ω aperto di R^k, μ misura in $\mathcal{L}_\Omega$].

E' costituito dalle funzioni reali definite μ-quasi ovunque in Ω, ivi μ-misurabili e limitate se $p=+\infty$, di potenza p-esima μ-sommabile se $p\in[1,+\infty[$ (cfr. n.6 e n.7 pag. 45).

Le operazioni sono quelle usuali. Se si considerano funzioni complesse, si ottiene lo spazio $L^p(\Omega,\mu)$ complesso.

4. Lo spazio l_p [p reale ≥ 1].

E' costituito dalle successioni reali:

$$x=(x_1,x_2,\dots,x_n,\dots)$$

di potenza p-esima sommabile [per $p=2$, di quadrato sommabile], cioè tali che la serie:

$$|x_1|^p+|x_2|^p+\dots+|x_n|^p+\dots$$

sia convergente.

Le operazioni sono quelle usuali: se $x=(x_n)_{n\in N}$, $y=(y_n)_{n\in N}$, $\lambda\in R$, si pone:

$$x+y=(x_1+y_1,x_2+y_2,\dots,x_n+y_n,\dots),\quad \lambda x=(\lambda x_1,\lambda x_2,\dots,\lambda x_n,\dots).$$

Il fatto che $x\in l_p, y\in l_p \Rightarrow x+y\in l_p$ si stabilisce in base alla (1) di pag. 54.

La famiglia di spazi $(l_p)_{p\in[1,+\infty[}$ è crescente rispetto alla relazione di inclusione. Si ha cioè: $p_1<p_2\Rightarrow l_{p_1}\subset l_{p_2}$.

Considerando successioni complesse, si ottiene lo spazio l_p complesso.

5. Lo spazio l_∞.

E' costituito dalle successioni reali limitate, con le stesse operazioni precisate nell'es. precedente.

Ovviamente per ogni $p\geq 1$ lo spazio l_p è un sottospazio di l_∞ [se la serie $\sum|x_n|^p$ è convergente, la successione $(|x_n|^p)_{n\in N}$ tende a zero, e quindi è limitata: tale è allora anche la successione $(x_n)_{n\in N}$].

Naturalmente si può considerare anche lo spazio l_∞ complesso.

6. L'insieme delle successioni reali [complesse], munito delle usuali operazioni di addizione e di moltiplicazione per uno scalare, è uno spazio vettoriale. Di questo sono ovviamente sottospazi gli l_p con $p\geq 1$ e l_∞.

7. L'insieme delle funzioni reali [complesse] di k variabili reali, definite in un sottoinsieme A di R^k, si può munire delle usuali operazioni di addizione e di moltiplicazione per uno scalare, ed acquista così la struttura di spazio vettoriale.

Se A è misurabile, invece delle funzioni "definite in A" si possono considerare le funzioni definite quasi ovunque in A (che sono classi di funzioni). Nel caso $A=\Omega$ aperto lo spazio vettoriale che si ottiene contiene $L^p(\Omega)$ e $L^\infty(\Omega)$, e questi sono sottospazi di esso.

8. L'insieme dei polinomi nel campo reale [complesso], munito delle usuali operazioni di addizione e di moltiplicazione per uno scalare, è uno spazio vettoriale che denotiamo con $\mathcal{P}$. Il sottospazio di $\mathcal{P}$ costituito dal polinomio nullo e dai polinomi di grado $\leq n$ sarà denotato con $\mathcal{P}_n$.

L'insieme costituito dai polinomi $p_0(t)=1$, $p_1(t)=t$, $p_2(t)=t^2$, ..., $p_n(t)=t^n$ è una base dello spazio vettoriale $\mathcal{P}_n$, sicché questo ha dimensione finita, uguale ad $n+1$.

Invece $\mathcal{P}$ ha ovviamente dimensione infinita, e così pure tutti gli spazi dei precedenti esempi 1, ..., 7 hanno dimensione infinita.

9. Lo spazio $C^0(T)$ $[T \subseteq R^k]$.

E' costituito dalle funzioni reali continue in T se T è compatto, continue e limitate in T se T non è compatto(1). Si tratta di un sottospazio di quello considerato nel precedente esempio 7 (caso funzioni reali).

10. Lo spazio $C^m(T)$ [m intero ≥ 1; T pseudodominio di R^k].

Col termine <u>pseudodominio</u> intendiamo un sottoinsieme di R^k che si ottiene da un aperto aggregando ad esso una parte (eventualmente vuota) della sua frontiera (anche tutta la frontiera).

Un insieme siffatto è caratterizzato dalla proprietà che ogni punto della frontiera è di accumulazione per l'interno.

Con $C^m(T)$ si denota il sottospazio di $C^0(T)$ costituito dagli elementi di $C^0(T)$ dotati in T di tutte le derivate fino a quelle di ordine m, tali derivate appartenenti anch'esse a $C^0(T)$.

Se $k>1$, detta $(\tau_1, \ldots, \tau_k)$ una k-pla di interi non negativi (chiamata <u>multiindice</u>), per indicare la derivata di un elemento $x=x(t)$ di $C^m(T)$, $t=(t_1, \ldots, t_k)$, eseguita τ_1 volte rispetto a t_1, ..., τ_k volte rispetto a t_k, si denota con τ il multiindice $(\tau_1, \ldots, \tau_k)$, con $|\tau|$ la somma $\tau_1 + \cdots + \tau_k$ (chiamata <u>norma</u> del multiindice), e si usa la notazione sintetica:

$$D^\tau x$$

che noi leggiamo, sia pure impropriamente, "derivata τesima" di x. Si tratta di una delle derivate parziali di ordine $|\tau|$. Naturalmente per $|\tau|=0$ si intende $D^\tau x = x$.

11. Lo spazio $C^\infty(T)$ [T pseudodominio di R^k].

E' il sottospazio di $C^0(T)$ costituito dalle funzioni $x \in C^0(T)$ appartenenti a $C^m(T)$ qualunque sia $m \in N$. E' l'intersezione di tutti gli spazi $C^m(T)$ con $m \in N$.

12. Lo spazio $C^{0,\lambda}(T)$ $[0<\lambda\leq 1;\ T \subseteq R^k]$.

E' il sottospazio di $C^0(T)$ costituito dalle funzioni $x \in C^0(T)$ che sono <u>hölderiane in T con esponente λ</u>. Tale locuzione significa che esiste $h \in R$ tale da risultare:

$$|x(t')-x(t'')| \leq h\,|t'-t''|^\lambda \qquad \forall t', t'' \in T.$$

Questa proprietà di $x(t)$ viene detta <u>condizione di Hölder</u> [<u>di Lipschitz</u> se $\lambda=1$].

Il più piccolo dei numeri h per cui sussiste la precedente disuguaglianza, ossia il numero:

$$\sup_{t', t'' \in T,\ t' \neq t''} \frac{|x(t')-x(t'')|}{|t'-t''|^\lambda}$$

(1) Per la topologia che sarà associata allo spazio occorre che le funzioni siano limitate. Il motivo sarà chiaro nel seguito.

si chiama il coefficiente di Hölder [di Lipschitz se $\lambda = 1$] della funzione $x = x(t)$ relativo all'esponente λ, e si denota col simbolo:

$$[x]_\lambda .$$

Naturalmente il riferimento a Hölder nelle precedenti locuzioni si adopera qualunque sia $\lambda \in]0,1]$, e quindi non esclude che sia $\lambda = 1$ anche se in tal caso si preferisce far riferimento a Lipschitz.

E' facile verificare che $\lambda_1 < \lambda_2 \Rightarrow C^{0,\lambda_1} \supset C^{0,\lambda_2}$: pertanto la famiglia di spazi $(C^{0,\lambda}(T))_{\lambda \in]0,1]}$ è decrescente rispetto all'inclusione.

In opportune ipotesi su T [ad es. se T è convesso, sicché ad ogni coppia di suoi punti è applicabile il teorema di Lagrange], lo spazio $C^1(T)$ è contenuto in $C^{0,1}(T)$, e quindi anche in $C^{0,\lambda}(T)$ $\forall \lambda \in]0,1[$.

Invitiamo il lettore a soffermarsi sull'argomento in Appendice a pag. 330.

13. Lo spazio $C^{m,\lambda}(T)$ [m intero ≥ 1 ; $0 < \lambda \leq 1$; T pseudopolinomio di R^N].

E' il sottospazio di $C^m(T)$ costituito dalle funzioni $x \in C^m(T)$ le cui derivate di ordine $\leq m$ appartengono a $C^{0,\lambda}(T)$.

Notiamo che in opportune ipotesi su T, ad esempio se T è convesso, nella precedente definizione basta supporre che appartengano a $C^{0,\lambda}(T)$ le derivate di ordine m, giacché quelle di ordine minore di m sono necessariamente hölderiane con qualunque esponente [in quanto appartengono a $C^1(T)$ e quindi sono lipschitziane].

14. Lo spazio $C_0^0(\Omega)$ [Ω aperto di R^N].

E' il sottospazio di $C^0(\Omega)$ costituito dalle funzioni $x \in C^0(\Omega)$ che sono a supporto compatto in Ω.

Per precisare il significato di tale locuzione, per ogni $x \in C^0(\Omega)$ consideriamo il sottoinsieme di Ω:

$$\Omega' = \{ t \in \Omega : x(t) \neq 0 \} .$$

La chiusura di Ω', che è contenuta in $\overline{\Omega}$, si chiama il supporto della funzione $x(t)$, e si denota col simbolo $\operatorname{supp} x$. Quando il supporto di x è contenuto in Ω ed è limitato (quindi compatto), si dice che la funzione $x = x(t)$ è a supporti compatto in Ω (1).

Se $x \in C_0^0(\Omega)$, per ogni $t_0 \in \Omega - \operatorname{supp} x$ risulta $x(t_0) = 0$, sicché la funzione $x(t)$ è identicamente nulla in $\Omega - \operatorname{supp} x$. Notiamo che, essendo continua, la $x(t)$ si annulla anche nei punti frontiera dell'insieme $\operatorname{supp} x$, ma per ciascuno di tali punti non esiste un intorno in cui $x(t)$ sia identicamente nulla.

Se $x \in C_0^0(\Omega)$, ovviamente la funzione $y(t)$ definita ponendo:

$$y(t) = \begin{cases} x(t) & \text{se } t \in \operatorname{supp} x \\ 0 & \text{se } t \notin \operatorname{supp} x \end{cases}$$

è un prolungamento di $x(t)$ continuo in tutto R^N. Avvertiamo che spesso tale prolungamento viene indicato con lo stesso simbolo $x(t)$: in altri termini, spesso una funzione continua $x(t)$ a supporto compatto si considera "automaticamente" definita in tutto lo spazio e nulla in $R^N - \operatorname{supp} x$.

15. Lo spazio $C_0^m(\Omega)$ [m intero ≥ 1 ; Ω aperto di R^N] (2).

E' il sottospazio di $C_0^0(\Omega)$ costituito dagli elementi di $C_0^0(\Omega)$ che appartengono anche a $C^m(\Omega)$.

---

(1) Il fatto che l'insieme degli elementi di $C^0(\Omega)$ a supporto compatto in Ω sia un sottospazio di $C^0(\Omega)$ consegue dalle ovvie relazioni:

$$\operatorname{supp}(x_1 + x_2) \subseteq (\operatorname{supp} x_1) \cup (\operatorname{supp} x_2) , \quad \operatorname{supp}(\lambda x) \subseteq \operatorname{supp} x \quad \forall \lambda \in R .$$

(2) Avvertiamo che alcuni autori in luogo di $C_0^m(\Omega)$ adoperano il simbolo $C_c^m(\Omega)$ richiamando con l'indice c la parola "compatto" (naturalmente anche per $m = 0$).

Ricordiamo che, se u, v sono elementi di $C^1(T)$, per la validità della formula di integrazione per parti:

$$\int_T (u D_i v)\,dt = \int_{\partial T} uv(\underline{e}_i \cdot \underline{n})\,d\sigma - \int_T (v D_i u)\,dt \qquad \left[t=(t_1,\dots t_n),\ D_i = \frac{\partial}{\partial t_i}\right]$$

dove $\underline{n}$ è il versore della normale esterna a ∂T ed $\underline{e}_i$ è il versore dell'asse t_i, è necessario che la frontiera di T soddisfi ad opportune condizioni (1). Però se u e v sono elementi di $C_0^1(\Omega)$, considerando le due funzioni prolungate a 0 in tutto lo spazio ed applicando la formula ad un dominio T che soddisfi alle menzionate condizioni e che contenga il supporto di u, si ottiene l'uguaglianza:

$$\int_T (u D_i v)\,dt = -\int_T (v D_i u)\,dt$$

nella quale, essendo $u=0$ e $D_i u = 0$ in T-suppu, l'insieme di integrazione può essere sostituito con suppu, e quindi anche con Ω [giacché in Ω-suppu si ha $u=0$, $D_i u = 0$]. Ne segue la formula:

$$(1) \qquad \int_\Omega (u D_i v)\,dt = -\int_\Omega (v D_i u)\,dt,$$

che sussiste senza alcuna condizione sulla frontiera di Ω. Si può anzi dimostrare che la (1) sussiste anche se una sola delle funzioni $u, v \in C^1(\Omega)$ è a supporto compatto in Ω.

Applicando reiteratamente la (1) si deduce che <u>per ogni $u \in C_0^m(\Omega)$ e per ogni $v \in C^m(\Omega)$, qualunque sia il multiindice α tale che $|\alpha| \le m$, risulta:</u>

$$(2) \qquad \int_\Omega (u D^\alpha v)\,dt = (-1)^{|\alpha|} \int_\Omega (v D^\alpha u)\,dt.$$

Nella (2), Ω è un aperto qualunque di R^k.

16. Lo spazio $C_0^\infty(\Omega)$ [Ω aperto di R^k].

E' il sottospazio di $C_0^0(\Omega)$ costituito dagli elementi di $C_0^0(\Omega)$ che appartengono anche a $C^\infty(\Omega)$.

17. Lo spazio $L^p_{loc}(\Omega)$ [p reale ≥ 1; Ω aperto di R^k].

E' costituito dalle funzioni $x(t)$ definite quasi ovunque in Ω tali che la potenza $|x(t)|^p$ sia "localmente sommabile", cioè misurabile in Ω e sommabile in ogni compatto contenuto in Ω.

E' utile osservare che, fissato $y \in C_0^0(\Omega)$, qualunque sia $x \in L^1_{loc}(\Omega)$ ha significato l'integrale:

$$\int_\Omega x(t) y(t)\,dt$$

giacché nell'insieme compatto suppy $\subset \Omega$ la funzione $x(t)$ è sommabile quindi è ivi sommabile anche la funzione $x(t)y(t)$ [essendo $y(t)$ continua], e d'altra parte $x(t)y(t)$ è identicamente nulla in Ω-suppy.

---

(1) Ad es. per $k=2$ la formula si stabilisce nell'ipotesi che T sia un dominio regolare, ossia un dominio limitato ad uno o più contorni ciascuno dei quali sia una curva generalmente regolare.

Osserviamo pure che per ogni $p \geq 1$ si ha $L^p(\Omega) \subset L^1_{loc}(\Omega)$: infatti se $u \in L^p(\Omega)$, per ogni compatto $K \subset \Omega$ si ha anche $u \in L^p(K)$ e quindi, essendo K limitato, $u \in L^1(K)$.

18. Derivate deboli. Lo spazio $W^{m,p}(\Omega)$ [$m \in \mathbb{N}$; $p \geq 1$; Ω aperto di $\mathbb{R}^n$].

Supponiamo in un primo tempo che Ω sia un intervallo aperto di $\mathbb{R}$, e consideriamo una funzione $u \in C^m(\Omega)$.

Per ogni funzione $\varphi \in C_0^\infty(\Omega)$ consideriamo l'integrale:

$$\int_\Omega u^{(m)} \varphi \, dt$$

ed applichiamo m volte la regola di integrazione per parti. Tenuto conto che $\varphi, \varphi', \dots, \varphi^{(m-1)}$ sono nulle negli estremi di Ω, si ottiene l'uguaglianza:

$$\int_\Omega u^{(m)} \varphi \, dt = (-1)^m \int_\Omega u \varphi^{(m)} dt .$$

Pertanto la derivata m^{ma} di u è un elemento w di $C^0(\Omega)$ tale da risultare:

$$(1) \qquad \int_\Omega w \varphi \, dt = (-1)^m \int_\Omega u \varphi^{(m)} dt \quad \forall \varphi \in C_0^\infty(\Omega).$$

Supponiamo ora che, viceversa, w sia un elemento di $C^0(\Omega)$ soddisfacente alla (1). Applicando al secondo membro m volte la regola di integrazione per parti, si ottiene:

$$\int_\Omega w \varphi \, dt = \int_\Omega u^{(m)} \varphi \, dt$$

e quindi risulta:

$$(2) \qquad \int_\Omega [w - u^{(m)}] \varphi \, dt = 0 \quad \forall \varphi \in C_0^\infty(\Omega) .$$

Si dimostra che la (2) implica $w - u^{(m)} = 0$, e quindi per le funzioni $u \in C^m(\Omega)$ sussiste l'equivalenza:

$$(3) \qquad w = u^{(m)} \iff \left[\int_\Omega w \varphi \, dt = (-1)^m \int_\Omega u \varphi^{(m)} dt \quad \forall \varphi \in C_0^\infty(\Omega) \right]$$

La (3) rappresenta una caratterizzazione della derivata m^{ma} di una funzione $u \in C^m(\Omega)$: si può dire che <u>la derivata m^{ma} di u è l'elemento w di $C^0(\Omega)$ che soddisfa l'uguaglianza (1)</u>.

Notiamo che la (1), ossia il secondo membro della (3), ha significato anche se u e w sono sommabili in ogni compatto contenuto in Ω, cioè elementi dello spazio $L^1_{loc}(\Omega)$: pertanto la (3) consente di generalizzare il concetto di derivata m^{ma}, estendendo tale nozione agli elementi di $L^1_{loc}(\Omega)$.

Precisamente, assegnato un elemento u di $L^1_{loc}(\Omega)$, se esiste un elemento w di $L^1_{loc}(\Omega)$ tale da verificare la (1), si dice che la funzione u <u>è dotata di derivata m^{ma} debole</u>; in tale ipotesi si dimostra che w è l'unico elemento di $L^1_{loc}(\Omega)$ soddisfacente alla (1), e w si chiama <u>la derivata m^{ma} debole di u</u>, che si denota con $u^{(m)}$.

Dunque per ogni $u \in L^1_{loc}(\Omega)$ la derivata n^{ma} debole, quando esiste, è un elemento dello stesso spazio $L^1_{loc}(\Omega)$, ed è l'unico elemento $u^{(n)}$ di $L^1_{loc}(\Omega)$ tale che risulti:

$$\int_\Omega u^{(n)} \varphi \, dt = (-1)^n \int_\Omega u \varphi^{(n)} dt \quad \forall \varphi \in C_0^\infty(\Omega).$$

Ricordiamo che un elemento di $L^1_{loc}(\Omega)$ è una funzione definita quasi ovunque in Ω (ivi localmente sommabile), ossia [cfr. pag. 42] è una classe di funzioni a due a due uguali quasi ovunque. Ricordiamo pure che l'integrale di una funzione definita q.o. in Ω (ivi sommabile) è per definizione l'integrale di uno qualunque dei rappresentanti della classe.

Se $\tilde{u}$ è una funzione appartenente a $C^n(\Omega)$, aggregando ad $\tilde{u}$ tutte le funzioni uguali q.o. ad $\tilde{u}$ si ottiene un elemento di $L^1_{loc}(\Omega)$, diciamolo u: per brevità di linguaggio si suol dire (identificando u con $\tilde{u}$) che $\tilde{u}$ appartiene ad $L^1_{loc}(\Omega)$. Da tale consuetudine deriva quella di considerare $C^n(\Omega)$ come sottospazio di L^1_{loc}: perciò si può parlare di derivata debole di un elemento di $C^n(\Omega)$.

Ovviamente la predetta $\tilde{u}$ è dotata di derivata debole, e la derivata debole di $\tilde{u}$ è la classe $u^{(n)}$ di tutte le funzioni uguali q.o. ad $\tilde{u}^{(n)}$. La consuetudine dianzi esposta consente di dire brevemente che <u>la derivata n^{ma} in senso classico</u>, quando esiste, <u>è anche la derivata n^{ma} debole</u>. Naturalmente non viceversa: se un elemento u di $L^1_{loc}(\Omega)$ è dotato di derivata n^{ma} debole, non è detto che fra i rappresentanti della classe u ce ne sia uno appartenente a $C^n(\Omega)$.

È importante evidenziare che, mentre la derivata n^{ma} debole di $u \in L^1_{loc}(\Omega)$ (quando esiste) appartiene allo stesso spazio $L^1_{loc}(\Omega)$, la derivata n^{ma} in senso classico di un elemento di $C^n(\Omega)$ non è detto che appartenga allo stesso spazio $C^n(\Omega)$ [è un elemento di $C^0(\Omega)$].

L'esistenza della derivata debole si caratterizza mediante la nozione di assoluta continuità [pag. 47]. Infatti si dimostra che le seguenti proposizioni, per ogni $u \in L^1_{loc}(\Omega)$, sono equivalenti:

a) <u>u è dotato della derivata debole u'.</u>

b) <u>esiste un rappresentante $\tilde{u}$ di u che è assolutamente continuo in ogni intervallo compatto contenuto in Ω.</u>

Con riferimento alla b), la funzione $\tilde{u}$ è derivabile quasi ovunque in Ω [teor. 4 pag. 47 e teor. 3 pag. 46], e la sua derivata $\tilde{u}'$ è un rappresentante della derivata debole u'.

Dalla equivalenza appena enunciata consegue che un elemento di $L^1_{loc}(\Omega)$, avente un rappresentante $\tilde{u} \in C^{n-1}(\Omega)$, è dotato della derivata debole $u^{(n)}$ se e solo se $\tilde{u}^{(n-1)}$ è assolutamente continuo in ogni intervallo compatto contenuto in Ω.

Si dimostra che per ogni $n > 1$, se u è dotata della derivata n^{ma} debole, u è anche dotata delle derivate deboli di ordine minore di n.

In base alla nozione di derivata debole si può intendere in un senso più generale il problema espresso da un'equazione differenziale. Ad esempio l'equazione differenziale:

$$u^{(n)} = f, \tag{4}$$

che è ambientata nello spazio $C^{(n)}$, si può ambientare nello spazio "più ampio" $L^1_{loc}(\Omega)$ considerando il seguente problema. Assegnata una funzione $f \in L^1_{loc}(\Omega)$, ricercare le funzioni $u \in L^1_{loc}(\Omega)$ dotate di derivata n^{ma} debole tale da verificare la (4). Ogni eventuale soluzione di tale problema si chiama una <u>soluzione debole</u>, o <u>soluzione in senso debole</u>, dell'equazione differenziale (4).

* * *

La nozione di derivata debole si introduce analogamente per le funzioni di più variabili. Sia Ω un aperto di R^k, e sia $\alpha = (\alpha_1, \dots, \alpha_k)$ un multiindice [n.10 pag. 56]. Assegnato un elemento u dello spazio $L^1_{loc}(\Omega)$, se esiste un elemento w dello stesso spazio tale da risultare:

$$(5) \qquad \int_\Omega w\varphi\,dt = (-1)^{|\alpha|} \int_\Omega u D^\alpha \varphi\,dt \quad \forall \varphi \in C_0^\infty(\Omega),$$

si dice che u è <u>dotato della derivata debole $D^\alpha u$</u>; l'elemento w di $L^1_{loc}(\Omega)$, univocamente determinato, si chiama <u>la derivata debole di u di multiindice α (di ordine $|\alpha|$)</u> e si denota con $D^\alpha u$. Pertanto $D^\alpha u$, se esiste, è l'elemento di $L^1_{loc}(\Omega)$ tale da risultare:

$$(6) \qquad \int_\Omega \varphi D^\alpha u\,dt = (-1)^{|\alpha|} \int_\Omega u D^\alpha \varphi\,dt \quad \forall \varphi \in C_0^\infty(\Omega).$$

Spesso nel linguaggio corrente l'attributo "debole" viene sottinteso, dato che non possono sorgere equivoci. Se l'elemento u di $L^1_{loc}(\Omega)$ ha un rappresentante di classe $C^m(\Omega)$, allora u è dotato della derivata debole $D^\alpha u$ qualunque sia il multiindice α con $|\alpha| \le m$, e la derivata debole $D^\alpha u$ ha un rappresentante che coincide con l'analoga derivata in senso classico.

Contrariamente a quanto accade per $k=1$, se $\Omega \subseteq R^k$ con $k>1$ dall'esistenza delle derivate deboli di un certo ordine $m \ge 2$ non consegue l'esistenza delle derivate deboli di ordine minore di m.

* * *

Come si è visto all'inizio di pag. 59, per ogni $p \ge 1$ lo spazio $L^p(\Omega)$, con $\Omega \subseteq R^k$, è contenuto in $L^1_{loc}(\Omega)$: perciò in particolare si può parlare di derivate deboli anche per gli elementi di $L^p(\Omega)$. Se un elemento u di $L^p(\Omega)$ è dotato della derivata debole $D^\alpha u$, non è detto che questa, elemento di $L^1_{loc}(\Omega)$, appartenga a sua volta ad $L^p(\Omega)$.

Orbene, lo spazio vettoriale costituito dalle funzioni $u \in L^p(\Omega)$ dotate di tutte le derivate deboli di ordine $\le m$, che appartengano anch'esse ad $L^p(\Omega)$, si denota con

$$W^{m,p}(\Omega).$$

Per $p=2$, in luogo di $W^{m,2}$ si scrive semplicemente

$$W^m(\Omega).$$

Nelle precedenti notazioni la lettera W richiama la parola weak (debole).

Se l'aperto Ω di R^k ha misura finita, <u>lo spazio vettoriale $C^m(\Omega)$ è contenuto in $W^{m,p}(\Omega)$</u> qualunque sia $p \ge 1$: occorre precisare tale asserzione, che è efficace ma difettosa, e si adopera per brevità di linguaggio.

Sia $\tilde{u}$ una funzione appartenente a $C^m(\Omega)$, e sia u la classe di funzioni che si ottiene aggregando ad $\tilde{u}$ tutte le funzioni q.o. uguali ad $\tilde{u}$; per ogni multiindice α tale che $|\alpha| \leq m$ si consideri la classe di funzioni che si ottiene aggregando a $D^\alpha \tilde{u}$ tutte le funzioni q.o. uguali a $D^\alpha \tilde{u}$: tale classe è la derivata debole $D^\alpha u$ (abbiamo già detto brevemente che una derivata in senso classico è anche derivata debole).

Se Ω ha misura finita la funzione $D^\alpha \tilde{u}$ (che è misurabile e limitata in quanto $\tilde{u} \in C^m(\Omega)$) è sommabile con qualunque esponente $p \geq 1$, e perciò $D^\alpha u$ è un elemento di $L^p(\Omega)$. Tale discorso vale anche per $|\alpha| = 0$, ossia per la funzione $\tilde{u}$. Identificando $\tilde{u}$ con u e $D^\alpha \tilde{u}$ con $D^\alpha u$, si può dire che $\tilde{u}$ e $D^\alpha \tilde{u}$ appartengono ad $L^p(\Omega)$. Con tali identificazioni si può dire che $\tilde{u}$ e tutte le sue derivate di ordine $\leq m$ appartengono ad $L^p(\Omega)$, e quindi che $\tilde{u}$ è un elemento di $W^{m,p}(\Omega)$.

Analogamente, se un dato elemento u di $W^{m,p}(\Omega)$ ha un rappresentante di classe $C^m(\Omega)$, si suol dire brevemente che u appartiene a $C^m(\Omega)$.

ESEMPIO 1. La funzione $\tilde{u}$ definita ponendo $\tilde{u}(x) = |x|$ $\forall x \in R$ è derivabile quasi ovunque in R, essendo $\tilde{u}'(x) = 1$ per $x > 0$ e $\tilde{u}'(x) = -1$ per $x < 0$.

Siano u l'elemento di $L^1_{loc}(R)$ che ha $\tilde{u}$ come rappresentante, e w l'elemento di $L^1_{loc}(R)$ che ha $\tilde{u}'$ come rappresentante.

Per ogni $\varphi \in C_0^\infty(R)$ si ha, tenuto conto che $\tilde{u}(0) = 0$ e che φ è a supporto compatto in R (per cui $\varphi(t) \to 0$ per $t \to -\infty$ e per $t \to +\infty$):

$$\int_{-\infty}^{0} u\varphi' dt = \int_{-\infty}^{0} \tilde{u}\varphi' dt = [\tilde{u}\varphi]_{-\infty}^{0} - \int_{-\infty}^{0} \tilde{u}'\varphi\, dt = -\int_{-\infty}^{0} w\varphi\, dt \,, \quad \int_{0}^{+\infty} u\varphi' dt = [\tilde{u}\varphi]_{0}^{+\infty} - \int_{0}^{+\infty} \tilde{u}'\varphi\, dt = -\int_{0}^{+\infty} w\varphi$$

e quindi:

$$\int_{-\infty}^{+\infty} u\varphi' dt = -\left(\int_{-\infty}^{0} w\varphi\, dt + \int_{0}^{+\infty} w\varphi\, dt \right) = -\int_{-\infty}^{+\infty} w\varphi\, dt \,.$$

Pertanto, essendo:

$$\int_R w\varphi\, dt = -\int_R u\varphi' dt \qquad \forall \varphi \in C_0^\infty(R) \tag{7}$$

si conclude che w è la derivata debole di u.

ESEMPIO 2. La funzione $\tilde{u}$ definita ponendo $\tilde{u}(x) = 0$ $\forall x < 0$, $\tilde{u}(x) = 1$ $\forall x \geq 0$, ossia la funzione scalino, è derivabile q.o. in R essendo $\tilde{u}'(x) = 0$ $\forall x \in R - \{0\}$.

Siano u l'elemento di $L^1_{loc}(R)$ che ha $\tilde{u}$ come rappresentante, e w l'elemento nullo dello spazio $L^1_{loc}(R)$: la (7) non sussiste, perché il primo membro è uguale a 0 e il secondo membro, stante la definizione di $\tilde{u}$, è uguale a $\varphi(0)$ [basta pensare ad un elemento φ di $C_0^\infty(R)$ che non si annulli nel punto 0].

Peraltro, in base all'equivalenza tra le proposizioni a) e b) di pag. 60 è ovvio che u non è dotata di derivata debole: negli intervalli compatti a cui il punto 0 sia interno la $\tilde{u}$ non è continua, e quindi non è assolutamente continua.

2. Operatori lineari . Funzionali. Forme bilineari, forme quadratiche.

Siano S, S' spazi vettoriali . Ogni funzione $A : S \to S'$ si chiama un operatore di S in S' ;l'operatore A si chiama un operatore lineare (o una trasformazione lineare), quando risulta :

$$(1) \quad A(x+y) = A(x) + A(y) \quad , \quad A(\lambda x) = \lambda A(x)$$

qualunque siano $x, y \in S$ e $\lambda \in K$ (1).

Un'applicazione biunivoca di S su S' che sia un operatore lineare si chiama un isomorfismo (algebrico) di S su S'. Se esiste un isomorfismo f di S su S', si dice che gli spazi vettoriali S ed S' sono (algebricamente) isomorfi; in tal caso, identificando ciascun vettore x di S con il vettore $f(x)$ di S', i due spazi vettoriali S ed S' si identificano (mediante f).

Quando non vi è possibilità di equivoci, per ogni $x \in S$ l'elemento $A(x)$ di S' si denota semplicemente con Ax . Se A è definito in un sottospazio H di S , tale sottospazio si chiama il dominio di A , e si denota con D_A ; anche se $H \subset S$, si dice ancora che A è un operatore lineare di S in S' aggiungendo la precisazione di dominio D_A : tale precisazione si omette quando $D_A = S$. .

Qualunque sia l'operatore lineare A di S in S' , e qualunque sia H sottospazio di S , evidentemente $A(H)$ è un sottospazio di S';evidente è altresì che , qualunque sia H' sottospazio di S' , l'insieme :

$$A^{-1}(H') = \{x \in S : Ax \in H'\}$$

è un sottospazio di S . In particolare ($H' = \{0\}$) , il sottospazio di S :

$$\{x \in S : Ax = 0\}$$

si chiama il nucleo dell'operatore lineare A , e si denota con

$$\operatorname{Ker} A \, .$$

Si verifica facilmente che, qualunque siano $S_0 \subseteq S$, $S'_0 \subseteq S'$ si ha:

$$\begin{array}{ccc} S_0 \text{ convesso} & \Longrightarrow & A(S_0) \text{ convesso} \\ S'_0 \text{ convesso} & \Longrightarrow & A^{-1}(S'_0) \text{ convesso} \end{array} .$$

L'operatore lineare di S in S :

$$I : x \in S \to x$$

(1) Il verificarsi delle (1) si può esprimere efficacemente dicendo che l'operatore A "conserva le operazioni" degli spazi : infatti il corrispondente della somma di due vettori coincide con la somma dei loro corrispondenti, e il corrispondente del prodotto di un vettore per il numero λ coincide col prodotto di λ per il corrispondente del vettore .

si chiama operatore identico di S in S, o identità in S.

Qualunque sia H sottospazio vettoriale di S, l'operatore lineare J che ad ogni $x \in H$ fa corrispondere x medesimo considerato come elemento di S, si chiama immersione (o iniezione) di H in S (1).

Ogni funzione F definita in un qualunque insieme S e a valori nell'insieme dei numeri reali, si chiama un funzionale definito in S (2). Nell'ipotesi che S sia uno spazio vettoriale, un funzionale F, definito in un sottoinsieme convesso S_0 di S, si dice convesso quando risulta:

$$(2) \qquad F(\lambda x+(1-\lambda)y) \leq \lambda F(x)+(1-\lambda)F(y) \qquad \forall x,y \in S_0 \text{ e } \forall \lambda \in [0,1];$$

si dice che F é strettamente convesso quando la (2) sussiste con $<$ in luogo di $\leq$ e $\lambda \in]0,1[$.

Il verificarsi della (2) equivale al fatto che l'insieme (epigrafo di f):

$$\{(x,t) \in S_0 \times R : t \geq F(x)\}$$

é una parte convessa dello spazio vettoriale $S \times R$.

Naturalmente, un funzionale lineare definito in S é un operatore lineare di S in R. I funzionali lineari sono particolari funzionali convessi.

Consideriamo ora un vettore x_0 di S e un sottospazio vettoriale H di S, sicché il sottoinsieme di S:

$$(3) \qquad S_0 = x_0 + H$$

é una varietà lineare di S (passante per x_0). Se H ha dimensione 1, ossia é del tipo $H = Kx_1$ con $x_1 \in S-\{0\}$, la varietà lineare (3) é una retta di S, e siccome é l'insieme degli $x \in S$ del tipo $x = x_0 + tx_1$ con $t \in K$, si dice che S_0 é la retta di equazione $x = x_0 + tx_1$.

Orbene, la varietà lineare S_0 é un iperpiano di S se e solo se H é il nucleo di un funzionale lineare [funzionale lineare complesso se S é sul campo complesso] non identicamente nullo; ossia, se e solo se esiste un funzionale lineare u, non identicamente nullo, tale che $H = \ker u$. Se $H = \ker u$, posto $u(x_0) = \alpha$ si ha:

$$S_0 = x_0 + \ker u = \{x \in S : u(x) = \alpha\},$$

e perciò si dice che S_0 é l'iperpiano di equazione $u(x) = \alpha$.

Dim. Supponiamo che H sia supplementare alla retta Kx_1, sicché S é somma diretta di H e di Kx_1: per ogni $x \in S$ esistono allora, univocamente determinati, $h \in H$ e $t \in K$ tali che $x = h + tx_1$. Detto u il funzionale che ad ogni $x \in S$ associa il predetto numero t, ovviamente u é lineare, e risulta $H = \ker u$.

Viceversa, sia $H = \ker u$, con u funzionale lineare definito in S non identicamente nullo, e sia $x_1 \in S$ tale che $u(x_1) \neq 0$. Proviamo che H é supplementare alla retta Kx_1. Che sia $S = H + Kx_1$ segue dal fatto che $\forall x \in S$, posto $t = u(x)/u(x_1)$, risulta $x = (x - tx_1) + tx_1$ e $u(x - tx_1) = 0$; d'altra parte, se $y \in H \cap Kx_1$ risulta $u(y) = 0$ e $y = tx_1$ con $t \in K$, e quindi $y = 0$ [essendo $0 = u(y) = u(tx_1) = tu(x_1)$ si ha $t = 0$ e quindi $y = tx_1 = 0$]: pertanto $H \cap Kx_1 = \{0\}$ e l'asserto é provato.

---

(1) Occorre distinguere tale applicazione dell'identità in H, che é un applicazione di H in H. Per rendersi conto della necessità di tale distinzione, basta pensare al caso in cui ad H é associata una topologia, e ad S un'altra topologia: dire che l'applicazione $J : x \in H \longrightarrow x \in S$ é continua, non equivale a dire che l'applicazione $I : x \in H \longrightarrow x \in H$ é continua.

(2) Tale locuzione si adopera anche se F é a valori nell'insieme dei numeri complessi: in tal caso, noi diremo che F é un funzionale complesso definito in S.

Esempi.

1. Fissato un numero reale c, la funzione u definita dalla relazione:

(1) $$u(x) = cx \qquad \forall x \in R$$

è un funzionale lineare definito in R.

2. Fissato un punto $c = (c_1, \dots, c_n) \in R^n$, la funzione u definita come segue:

(2) $$u(x_1, x_2, \dots, x_n) = c_1 x_1 + c_2 x_2 + \cdots + c_n x_n \qquad \forall (x_1, \dots, x_n) \in R^n$$

è un funzionale lineare definito in R^n. Notiamo che, denotando con xy il prodotto scalare dei vettori x ed y di R^n, posto $x = (x_1, \dots, x_n)$ la (2) si può scrivere:

(2') $$u(x) = cx \qquad \forall x \in R^n$$

e quindi assume la stessa forma della (1).

Si verifica subito che tutti i funzionali lineari definiti in R^n sono del tipo (2), cioè sono polinomi omogenei di primo grado in n variabili reali. Infatti se v è un qualunque funzionale lineare definito in R^n, detta $(e_1, \dots, e_n)$ la base usuale dello spazio vettoriale R^n, per ogni $x = (x_1, \dots, x_n) \in R^n$ si ha:

$$v(x) = v(x_1 e_1 + x_2 e_2 + \cdots + x_n e_n) = x_1 v(e_1) + x_2 v(e_2) + \cdots + x_n v(e_n)$$

da cui, posto $c_i = v(e_i)$, segue $v(x_1, \dots x_n) = c_1 x_1 + c_2 x_2 + \cdots + c_n x_n$.

Naturalmente quanto sopra vale anche se si sostituisce R^n con C^n: in luogo della (2') si avrà in tal caso $u(x) = x\bar{c}$, in quanto il prodotto scalare $x \cdot y$ dei vettori $x = (x_1, \dots, x_n)$ ed $y = (y_1, \dots, y_n)$ di C^n è $x_1 \bar{y}_1 + x_2 \bar{y}_2 + \cdots + x_n \bar{y}_n$.

3. Fissata una matrice di tipo (m, n) ad elementi reali:

(3) $$\begin{pmatrix} c_{11} & c_{12} & \dots & c_{1n} \\ c_{21} & c_{22} & \dots & c_{2n} \\ \dots & \dots & \dots & \dots \\ c_{m1} & c_{m2} & \dots & c_{mn} \end{pmatrix}$$

consideriamo la corrispondenza definita in R^n:

$$A : (x_1, x_2, \dots, x_n) \in R^n \longrightarrow \begin{pmatrix} c_{11} & c_{12} & \dots & c_{1n} \\ c_{21} & c_{22} & \dots & c_{2n} \\ \dots & \dots & \dots & \dots \\ c_{m1} & c_{m2} & \dots & c_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} .$$

Per ogni $(x_1, \dots, x_n) \in R^n$, il corrispondente $A(x_1, \dots, x_n)$ è la matrice ad una colonna (vettore colonna):

(4) $$\begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix}$$

con $y_1 = c_{11} x_1 + \cdots + c_{1n} x_n$, $\dots$, $y_m = c_{m1} x_1 + \cdots + c_{mn} x_n$. Convenendo di identificare ciascun punto $(y_1, \dots y_m)$ di R^m con il vettore colonna (4), si può dire che A è un operatore di R^n in R^m, ed è facile verificare che si tratta di un operatore lineare.

Notiamo che, se si indica con c la matrice (3), posto $x = (x_1, \dots, x_n)$ l'operatore A si può definire mediante l'uguaglianza:

(5) $$Ax = cx \qquad \forall x \in R^n$$

dove il secondo membro si intende, naturalmente, come il prodotto della matrice c per la matrice x. Così anche l'operatore lineare A di R^n in R^m assume la stessa forma del funzionale lineare (1) relativo al caso $n = m = 1$.

Si verifica subito che tutti gli operatori lineari di R^k in R^m sono del tipo dianzi considerato. Infatti, se B è un qualunque operatore lineare di R^k in R^m, detta $(e_1, \dots, e_k)$ la base usuale dello spazio vettoriale R^k, per ogni $x = (x_1, \dots, x_k) \in R^k$ si ha:

$$Bx = B(x_1 e_1 + \cdots + x_k e_k) = x_1 B e_1 + x_2 B e_2 + \cdots x_k B e_k ;$$

di conseguenza, posto:

$$Be_1 = \begin{pmatrix} c_{11} \\ c_{21} \\ \vdots \\ c_{m1} \end{pmatrix}, \quad Be_2 = \begin{pmatrix} c_{12} \\ c_{22} \\ \vdots \\ c_{m2} \end{pmatrix}, \dots, \quad Be_k = \begin{pmatrix} c_{1k} \\ c_{2k} \\ \vdots \\ c_{mk} \end{pmatrix}$$

risulta:

$$Bx = \begin{pmatrix} c_{11} x_1 \\ c_{21} x_1 \\ \vdots \\ c_{m1} x_1 \end{pmatrix} + \begin{pmatrix} c_{12} x_2 \\ c_{22} x_2 \\ \vdots \\ c_{m2} x_2 \end{pmatrix} + \cdots + \begin{pmatrix} c_{1k} x_k \\ c_{2k} x_k \\ \vdots \\ c_{mk} x_k \end{pmatrix} = \begin{pmatrix} c_{11} x_1 + c_{12} x_2 + \cdots + c_{1k} x_k \\ c_{21} x_1 + c_{22} x_2 + \cdots + c_{2k} x_k \\ \cdots\cdots\cdots\cdots \\ c_{m1} x_1 + c_{m2} x_2 + \cdots + c_{mk} x_k \end{pmatrix}$$

ossia:

$$Bx = \begin{pmatrix} c_{11} & c_{12} & \cdots & c_{1k} \\ c_{21} & c_{22} & \cdots & c_{2k} \\ \cdots & \cdots & \cdots & \cdots \\ c_{m1} & c_{m2} & \cdots & c_{mk} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_k \end{pmatrix} .$$

Dunque gli operatori lineari di R^k in R^m sono in corrispondenza biunivoca con le matrici di tipo (m,k) ad elementi reali. Tenuto conto anche della (5), nel linguaggio corrente si suole identificare l'operatore lineare A con la matrice (3), e si suol dire perciò che un operatore lineare di R^k in R^m "è una matrice" di tipo (m,k).

4. Gli spazi vettoriali che qui prenderemo in considerazione sono sul campo reale.

Sia S_k uno spazio vettoriale di dimensione finita k. Fissata una base $(e_1, \dots, e_k)$ di S_k, per ogni $x \in S_k$ e $\forall i \in (1, \dots, k)$ denotiamo con $p_i(x)$ la componente i^{ma} del vettore x rispetto alla base $(e_1, \dots, e_k)$. Fissato un punto $c = (c_1, \dots, c_k) \in R^k$, la funzione u definita ponendo:

$$(6) \qquad u(x) = c_1 p_1(x) + c_2 p_2(x) + \cdots + c_k p_k(x) \quad \forall x \in S_k$$

è un funzionale lineare definito in S_k. Si verifica subito che tutti i funzionali lineari definiti in S_k sono del tipo (6).

Detto S_m uno spazio vettoriale di dimensione finita m, sia poi $(i_1, \dots i_m)$ una base di S_m. Fissata la matrice (3) ad elementi reali, per ogni $x \in S_k$ resta determinata la matrice

$$(7) \qquad \begin{pmatrix} c_{11} p_1(x) + c_{12} p_2(x) + \cdots + c_{1k} p_k(x) \\ c_{21} p_1(x) + c_{22} p_2(x) + \cdots + c_{2k} p_k x \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ c_{m1} p_1(x) + c_{m2} p_2(x) + \cdots + c_{mk} p_k x \end{pmatrix} = \begin{pmatrix} c_{11} & c_{12} \dots c_{1k} \\ c_{21} & c_{22} \dots c_{2k} \\ \cdots & \cdots \cdots \\ c_{m1} & c_{m2} \dots c_{mk} \end{pmatrix} \begin{pmatrix} p_1(x) \\ p_2(x) \\ \vdots \\ p_k(x) \end{pmatrix} .$$

La corrispondenza A che ad ogni $x \in S_k$ associa il vettore di S_m le cui componenti rispetto alla base $(i_1, \dots, i_m)$ sono le righe della matrice (7), è ovviamente un operatore lineare di S_k in S_m. È facile verificare che tutti gli operatori lineari di S_k in S_m sono del tipo dianzi considerato.

Possiamo dunque concludere che, se S_k ed S_m sono spazi vettoriali a dimensione finita, di dimensioni rispettive k ed m, fissate una base di S_k ed una base di S_m gli operatori lineari di S_k in S_m sono in corrispondenza biunivoca con le matrici di tipo (m,k) ad elementi reali. Del resto tale conclusione si deduce subito dal precedente es. 3, rimarcando preventivamente che uno spazio vettoriale di dimensione finita h si può identificare, fissata una sua base, con lo spazio R^h.

Quanto sopra si può ripetere considerando spazi vettoriali sul campo complesso: naturalmente si sostituiranno i numeri reali con numeri complessi, ed R^k con C^k.

5. Fissato $z \in L^2(\Omega)$, essendo Ω un aperto di R^K, la corrispondenza:

$$(8) \qquad u : x \in L^2(\Omega) \to \int_\Omega x(t) z(t)\, dt$$

è un funzionale lineare definito in $L^2(\Omega)$. Più in generale, se p, p' sono numeri reali ≥ 1 tali che $\frac{1}{p} + \frac{1}{p'} = 1$, fissato $z \in L^{p'}(\Omega)$ la corrispondenza:

$$(9) \qquad u : x \in L^p(\Omega) \to \int_\Omega x(t) z(t)\, dt$$

è un funzionale lineare definito in $L^p(\Omega)$.

6. Fissato $z \in L^2(\Omega)$, con Ω aperto di R^K, la corrispondenza:

$$A : x \in L^2(\Omega) \to z \int_\Omega x(t) z(t)\, dt$$

è un operatore lineare di $L^2(\Omega)$ in $L^1(\Omega)$. Più in generale, se p, p' sono numeri reali ≥ 1 tali che $\frac{1}{p} + \frac{1}{p'} = 1$, fissati $z_0 \in L^{p'}(\Omega)$ e $z_1 \in L^p(\Omega)$, la corrispondenza:

$$(10) \qquad A : x \in L^p(\Omega) \to z_1 \int_\Omega x(t) z_0(t)\, dt$$

è un operatore lineare di $L^p(\Omega)$ in $L^p(\Omega)$. Detto S uno spazio vettoriale qualunque, e fissati $z_0 \in L^{p'}(\Omega)$, $z_1 \in S$, la (10) è un operatore lineare di $L^p(\Omega)$ in S.

7. Fissato $z \in L^1_{loc}(\Omega)$, la corrispondenza:

$$(11) \qquad u : x \in C_0^K(\Omega) \to \int_\Omega x(t) z(t)\, dt$$

è un funzionale lineare definito in $C_0^K(\Omega)$, e la corrispondenza:

$$(12) \qquad u : x \in C_0^\infty(\Omega) \to \int_\Omega x(t) z(t)\, dt$$

è un funzionale lineare definito in $C_0^\infty(\Omega)$. In particolare le (11) e (12) si possono considerare per ogni fissato $z \in L^1(\Omega)$.

8. Fissati $a_{ij} \in C^0(T)$, $b_i \in C^0(T)$, $c \in C^0(T)$ [$i = 1, \dots, K$; $j = 1, \dots, K$; T pseudodominio di R^K], la corrispondenza:

$$A : x \in C^2(T) \to \sum_{i,j}^{1,\dots,K} a_{ij} \frac{\partial^2 x}{\partial t_i \partial t_j} + \sum_{i=1}^{K} b_i \frac{\partial x}{\partial t_i} + c x$$

è un operatore lineare di $C^2(T)$ in $C^0(T)$.

9. Siano p, p' numeri reali tali che $1/p + 1/p' = 1$, e sia α un multiindice tale che $|\alpha| = m$. Fissato $z \in W^{m,p'}(\Omega)$, la corrispondenza:

$$u : x \in W^{m,p}(\Omega) \to \int_\Omega D^\alpha x(t)\, D^\alpha z(t)\, dt$$

è un funzionale lineare definito in $W^{m,p}(\Omega)$.

Supponiamo ora che S ed S' siano spazi vettoriali sul campo reale, e denotiamo con $b=b(x,y)$ un funzionale definito in $S\times S'$. Si dice che $b(x,y)$ e' un funzionale bilineare, o una forma bilineare, in $S\times S'$, quando e' lineare rispetto ad x per ogni $y\in S'$ e lineare rispetto ad y per ogni $x\in S$; cioe' quando per ogni fissato $y\in S'$ il funzionale definito in S:

$$x\in S \longrightarrow b(x,y) \tag{3}$$

e' lineare, e per ogni fissato $x\in S$ il funzionale definito in S':

$$y\in S' \longrightarrow b(x,y) \tag{4}$$

e' lineare.

Piu' generalmente, supponiamo che S ed S' siano spazi vettoriali sul campo K.

Si dice che $b(x,y)$ e' un funzionale sesquilineare, o una forma sesquilineare, in $S\times S'$, quando e' lineare rispetto ad x per ogni $y\in S'$, mentre rispetto ad y, per ogni fissato $x\in S$, conserva l'operazione di addizione, cioe' verifica la relazione:

$$b(x,y_1+y_2)=b(x,y_1)+b(x,y_2) \qquad \forall\, y_1,y_2\in S',$$

ma $b(x,\lambda y)$ e' uguale a $\overline{\lambda}b(x,y)$ anziche' a $\lambda b(x,y)$:

$$b(x,\lambda y)=\overline{\lambda}\,b(x,y) \qquad \forall\,\lambda\in K \text{ e } \forall y\in S' \quad (1).$$

Naturalmente nel caso $K=R$ le locuzioni "forma sesquilineare" e "forma bilineare" hanno il medesimo significato. Avvertiamo che alcuni Autori adoperano il termine "bilineare" anche quando $K=C$.

Le forme sesquilineari [bilineari] in $S\times S$, ossia con $S'=S$, talvolta si chiamano piu' brevemente forme sesquilineari [bilineari] in S.

La forma sesquilineare $b(x,y)$ in S si dice simmetrica (o hermitiana) quando risulta:

$$b(x,y)=\overline{b(y,x)}\,. \tag{5}$$

La forma sesquilineare $b(x,y)$ in S da' luogo al funzionale (funzionale complesso se $K=C$):

$$F:\ x\in S \longrightarrow b(x,x)$$

che si chiama una forma quadratica in $S\times S$, o brevemente una forma quadratica in S (generata da $b(x,y)$). Evidentemente risulta:

---

(1) Ad esempio, se $x\cdot y$ e' il prodotto scalare dei vettori $x=(x_1,\dots,x_K)$, $y=(y_1,\dots,y_K)$ di C^K, cioe' se:

$$x\cdot y=\sum_{i=1}^{K} x_i\,\overline{y_i}\,,$$

ponendo $b(x,y)=x\cdot y$ si definisce una forma sesquilineare in $C^K\times C^K$.

$$F(\lambda x) = |\lambda|^2 F(x) \qquad \forall x \in S \text{ e } \forall \lambda \in K .$$

Osserviamo che, se la forma sesquilineare $b(x,y)$ è simmetrica, la forma quadratica da essa generata assume soltanto valori reali, in quanto la (5) implica $b(x,x) = \overline{b(x,x)}$ $\forall x \in S$.

Esempi.

1. Fissata una matrice (c_{ij}) di tipo (k,m) ad elementi complessi, la funzione complessa definita come segue:

$$b(x,y) = \sum_{i=1}^{k} \sum_{j=1}^{m} c_{ij}\, x_i \overline{y}_j \qquad \forall x = (x_1,\dots,x_k) \in C^k \text{ e } \forall y = (y_1,\dots,y_m) \in C^m \tag{6}$$

è una forma sesquilineare in $C^k \times C^m$. Per $k=m$ la (6) è una forma sesquilineare in C^k, e genera la forma quadratica in C^k:

$$F(x) = \sum_{i,j}^{1\dots k} c_{ij}\, x_i \overline{x}_j \qquad \forall x = (x_1,\dots,x_k) \in C^k . \tag{7}$$

Si verifica subito che tutte le forme sesquilineari in $C^k \times C^m$ sono del tipo (6). Infatti se $b_1(x,y)$ è una forma sesquilineare in $C^k \times C^m$, siccome per ogni $y \in C^m$ la $b_1(x,y)$ è un funzionale lineare definito in C^k, la $b_1(x,y)$ è del tipo:

$$b_1(x,y) = \sum_{i=1}^{k} c_i(y)\, x_i$$

essendo le $c_i(y)$ funzioni complesse. Siccome poi $b_1(x,\overline{y})$ come funzione di y è un funzionale lineare definito in C^m, le funzioni $c_i(\overline{y})$ devono essere del tipo:

$$c_i(\overline{y}) = \sum_{j=1}^{m} c_{ij}\, y_j$$

e dalle ultime due relazioni consegue che $b_1(x,y) = \sum_{i=1}^{k} \sum_{j=1}^{m} c_{ij}\, x_i \overline{y}_j$.

Da quanto sopra si deduce che le forme quadratiche in C^k sono tutte del tipo (7), e che le forme quadratiche in R^k sono i polinomi omogenei di secondo grado in k variabili reali.

Notiamo che per $k=m$ la forma sesquilineare (6) è simmetrica quando risulta $c_{ij} = \overline{c}_{ji}$, cioè quando la matrice (c_{ij}) coincide con la trasposta della matrice coniugata (1). Per le forme bilineari in R^k ciò vuol dire che la matrice (c_{ij}) è simmetrica.

2. Fissata una matrice (a_{ij}) i cui elementi appartengano ad $L^\infty(\Omega)$, nonché una elemento c dello spazio $L^\infty(\Omega)$ [Ω aperto di R^k], la funzione:

$$b(x,y) = \sum_{i,j}^{1\dots k} \int_\Omega a_{ij}(t) \frac{\partial x}{\partial t_i} \frac{\partial y}{\partial t_j}\, dt + \int_\Omega c(t)\, x(t)\, y(t)\, dt \qquad \forall x,y \in W^1(\Omega)$$

è una forma bilineare in $W^1(\Omega)$, simmetrica se la matrice (a_{ij}) è simmetrica.

---

(1) In particolare ciò accade se $c_{ij} = \delta_{ij}$, essendo δ_{ij} il simbolo di Kronecker:

$$\delta_{ij} = \begin{cases} 0 & \text{se } i \neq j \\ 1 & \text{se } i = j . \end{cases}$$

In tal caso la (6) si riduce al prodotto scalare in C^k.

3. Lo spazio degli operatori lineari di Σ in Σ'. Duale algebrico di uno spazio vettoriale.

Siano Σ, Σ' spazi vettoriali. Col simbolo $\mathcal{L}(\Sigma,\Sigma')$ denotiamo l'insieme costituito dagli operatori lineari di Σ in Σ'.

Qualunque siano $A,B \in \mathcal{L}(\Sigma,\Sigma')$, l'operatore lineare di Σ in Σ':

$$x \in \Sigma \longrightarrow Ax+Bx$$

si denota con $A+B$; qualunque siano $\lambda \in K$ ed $A \in \mathcal{L}(\Sigma,\Sigma')$, l'operatore lineare di Σ in Σ':

$$x \in \Sigma \longrightarrow \lambda A x$$

si denota con λA. Abbiamo così introdotto una legge di composizione interna in $\mathcal{L}(\Sigma,\Sigma')$ e una legge di composizione esterna tra K ed $\mathcal{L}(\Sigma,\Sigma')$: è immediato constatare che la terna $(\mathcal{L}(\Sigma,\Sigma'),+,\cdot)$ è uno spazio vettoriale, che sarà denotato semplicemente con $\mathcal{L}(\Sigma,\Sigma')$.

Qualunque sia lo spazio vettoriale Σ sul campo reale, lo spazio $\mathcal{L}(\Sigma,R)$ si chiama <u>il duale algebrico</u> di Σ: pertanto il duale algebrico di Σ è lo spazio vettoriale costituito da tutti i funzionali lineari definiti in Σ.

Se Σ è uno spazio vettoriale sul campo complesso, per <u>duale algebrico</u> di Σ si intende non lo spazio $\mathcal{L}(\Sigma,C)$, ma lo spazio costituito dai funzionali complessi v tali che $\overline{v} \in \mathcal{L}(\Sigma,C)$, cioè tali che v sia il coniugato di un funzionale lineare [1].

Per denotare il duale algebrico dello spazio vettoriale Σ vengono adottati simboli diversi dai vari Autori: noi lo indicheremo con Σ^+.

L'operazione che ad ogni coppia $(u,x) \in \Sigma^+ \times \Sigma$ associa il valore del funzionale u nel punto x, si chiama <u>la dualità tra Σ e Σ^+</u>; quando il numero $u(x)$ (reale se Σ è reale, complesso se Σ è complesso) si interpreta come il risultato di detta operazione, tale numero si denota col simbolo:

$$\langle u,x \rangle \quad [2].$$

Si osservi che la dualità tra Σ e Σ^+:

$$(u,x) \in \Sigma^+ \times \Sigma \longrightarrow \langle u,x \rangle$$

è una forma sesquilineare [forma bilineare se Σ è reale] in $\Sigma^+ \times \Sigma$; cioè risulta, per definizione di u_1+u_2 e di λu

$$\langle u_1+u_2,x \rangle = \langle u_1,x \rangle + \langle u_2,x \rangle \,, \quad \langle \lambda u,x \rangle = \lambda \langle u,x \rangle \qquad \forall u_1,u_2,u \in \Sigma^+, x \in \Sigma, \lambda \in K$$

nonché, per le proprietà di $u \in \Sigma^+$:

$$\langle u,x_1+x_2 \rangle = \langle u,x_1 \rangle + \langle u,x_2 \rangle \,, \quad \langle u,\lambda x \rangle = \overline{\lambda} \langle u,x \rangle \qquad \forall x_1,x_2,x \in \Sigma \,, u \in \Sigma^+, \lambda \in K.$$

---

[1] Si osservi che $\overline{v} = u \in \mathcal{L}(\Sigma,C) \Longleftrightarrow v = \overline{u}$ con $u \in \mathcal{L}(\Sigma,C)$. Pertanto risulta $v(x_1+x_2) = v(x_1)+v(x_2)$ $\forall x_1,x_2 \in \Sigma$, nonché $v(\lambda x) = \overline{\lambda}\, v(x)$ $\forall x \in \Sigma$ e $\forall \lambda \in K$ in quanto si ha:

$$v(\lambda x) = \overline{\overline{v}(\lambda x)} = \overline{\lambda \overline{v}(x)} = \overline{\lambda}\, v(x).$$

[2] Avvertiamo che alcuni Autori adoperano il simbolo $\langle x,u \rangle$.

Esempi.

1. Per ogni $c=(c_1,\dots,c_k)\in R^k$ denotiamo con u_c il funzionale lineare definito in R^k ponendo:

$$(1)\qquad u_c(x)=c_1x_1+c_2x_2+\cdots+c_kx_k \qquad \forall x=(x_1,\dots,x_k)\in R^k .$$

Posto $\Sigma=R^k$, la corrispondenza:

$$f: c\in R^k \longrightarrow u_c\in \Sigma^+$$

e' un'applicazione biunivoca di R^k su Σ^+ [cfr. es. 2 pag. 64], e siccome e' ovviamente lineare la f e' un isomorfismo di R^k su Σ^+. Pertanto lo spazio vettoriale R^k e' isomorfo al suo duale, e si può dire che, a meno di isomorfismi, lo spazio vettoriale R^k coincide con il suo duale.

Analogamente si conclude che un qualunque spazio a dimensione finita e' isomorfo al suo duale, giacché, detta m la dimensione, il duale e' isomorfo allo spazio K^m [cfr. es. 4 pag. 65]. Pertanto si può dire che ogni spazio vettoriale a dimensione finita "coincide" con il suo duale.

Denotando con $x\cdot y$ il prodotto scalare dei vettori $x=(x_1,\dots,x_n)$ ed $y=(y_1,\dots,y_n)$ di R^k, la (1) si può scrivere:

$$\langle u_c, x\rangle = c\cdot x \qquad \forall x\in R^k$$

e quindi, identificando u_c con c $\forall c\in R^k$, si può dire che la dualità tra R^n e il suo duale [identificato con R^k] si riduce al prodotto scalare in R^k.

2. Dati p e p' due numeri reali >1 tali che $1/p+1/p'=1$, per ogni $z\in L^{p'}(\Omega)$ sia u_z il funzionale lineare definito in $L^p(\Omega)$:

$$(2)\qquad u_z : x\in L^p(\Omega)\longrightarrow \int_\Omega z(t)x(t)\,dt ,$$

che e' un elemento del duale algebrico di $L^p(\Omega)$.

Si dimostra che gli elementi del duale algebrico di $L^p(\Omega)$ non sono tutti del tipo (2), ossia che la corrispondenza $f: z\in L^{p'}(\Omega)\longrightarrow u_z\in (L^p(\Omega))^+$ non e' suriettiva. Peraltro, detto Λ il codominio di f [sottospazio di $(L^p(\Omega))^+$], la f e' un isomorfismo di $L^{p'}$ su Λ. Si può dire allora che il sottospazio Λ del duale algebrico di $L^p(\Omega)$ coincide, a meno di isomorfismi, con lo spazio vettoriale $L^{p'}(\Omega)$.

Notiamo che la (2) fornisce:

$$\langle u_z, x\rangle = \int_\Omega z(t)x(t)\,dt \qquad \forall x\in L^p(\Omega) \text{ e } \forall z\in L^{p'}(\Omega).$$

Vogliamo concludere questo numero segnalando una importante definizione, e a tale scopo premettiamo un esempio.

Per ogni $x=(x_1,\dots,x_k)\in R^k$ denotiamo con $F(x)$ il polinomio omogeneo di primo grado, di coefficienti $x_1,\dots,x_k$, nelle k variabili $t_1,\dots,t_k$:

$$(3)\qquad (t_1,\dots,t_k)\in R^k \longrightarrow x_1t_1+x_2t_2+\cdots+x_kt_k ,$$

che è un funzionale lineare definito in R^K e quindi è un elemento del duale algebrico di R^K. Nasce così la corrispondenza

$$F : R^K \longrightarrow (R^K)^+ \tag{4}$$

che ad ogni $x \in R^K$ associa il polinomio (3).

Qualunque siano i vettori $x=(x_1,\dots,x_K)$ ed $y=(y_1,\dots,y_K)$ di R^K si ha dunque, per ogni $t=(t_1,\dots,t_K) \in R^K$:

$$< F(x), t > = x_1t_1 + x_2t_2 + \cdots + x_Kt_K \quad , \quad < F(y), t > = y_1t_1 + y_2t_2 + \cdots + y_Kt_K$$

e quindi:

$$< F(x)-F(y), t > = (x_1-y_1)t_1 + (x_2-y_2)t_2 + \cdots + (x_K-y_K)t_K \, .$$

Ne segue, per $t_1=x_1-y_1$, $t_2=x_2-y_2$, ..., $t_K=x_K-y_K$, ossia per $t=x-y$:

$$< F(x)-F(y), x-y > = (x_1-y_1)^2 + (x_2-y_2)^2 + \cdots + (x_K-y_K)^2 ,$$

e quindi risulta:

$$< F(x)-F(y), x-y > \geq 0 \qquad \forall x,y \in R^K , \tag{5}$$

verificandosi l'uguaglianza se e solo se $x=y$.

Come abbiamo osservato nell'es. 1 di pag. 70, se si identifica R^K con il suo duale mediante la corrispondenza F, la dualità tra R^K e il suo duale si riduce al prodotto scalare in R^K, e per $K=1$ si riduce alla moltiplicazione in R. Con la predetta identificazione e per $K=1$ la (5) esprime allora che i numeri $F(x)-F(y)$, $x-y$ sono concordi, cioè che F è una funzione crescente [in senso stretto per quanto rilevato dopo la (5)].

Per tal motivo qualunque sia K si dice che la funzione (4) è un operatore strettamente monotono.

Orbene, detto Σ un qualunque spazio vettoriale sul campo reale, un'applicazione:

$$F : \Sigma \longrightarrow \Sigma^+$$

si chiama un operatore monotono quando risulta, qualunque siano $x,y \in \Sigma$:

$$< F(x)-F(y), x-y > \geq 0 \quad ; \tag{6}$$

se nella (6) l'uguaglianza si verifica solo per $x=y$, allora F si chiama un operatore strettamente monotono.

La teoria degli operatori monotoni ha avuto importanti sviluppi, sui quali peraltro non ci soffermeremo.

4. Prodotto di operatori. Operatori invertibili; inverso destro e inverso sinistro.

Siano: S, T, Σ spazi vettoriali, e

$$A: S \to T \quad , \quad B: T \to \Sigma$$

operatori lineari, il primo di dominio D_A, il secondo di dominio D_B.

Detto D il sottospazio di S:

$$(1) \qquad D = \{x \in S : x \in D_A \text{ e } Ax \in D_B\}$$

l'operatore lineare di S in Σ di dominio D:

$$x \in D \longrightarrow B(Ax) \in \Sigma$$

si chiama il prodotto degli operatori A e B, e si denota con BA.

Ogni volta che considereremo il prodotto di due operatori lineari A e B, riterremo sottintesa la condizione che il codominio di A e il dominio di B siano sottospazi dello stesso spazio vettoriale (1); quando sarà opportuno sottolineare che tale condizione è verificata, la esprimeremo concisamente dicendo che il prodotto BA "ha senso".

È facile verificare che se A è un operatore lineare di R^k in R^m, identificato con una matrice di tipo (m,k) [cfr. esempio 3 pag. 64], e B è un operatore lineare di R^m in R^p, identificato con una matrice di tipo (p,m), il prodotto BA degli operatori A e B si identifica con la matrice prodotto della matrice B per la matrice A. Si osservi che se $k \neq p$ ha senso l'operatore BA ma non ha senso l'operatore AB.

Da notare che, detti A e B due operatori lineari, se hanno senso entrambi gli operatori AB, BA, non risulta necessariamente $AB = BA$. Detto C un ulteriore operatore lineare, si ha:

$$C(BA) = (CB)A \quad , \quad C(A+B) = CA + CB \quad , \quad (A+B)C = AC + BC$$

ciascuna uguaglianza essendo valida purché abbiano senso entrambi i membri.

Osserviamo infine che, detta I l'identità in S e I' l'identità in T, qualunque sia l'operatore lineare $A: S \to T$ risulta $AI = I'A = A$.

Siano ora S, S' spazi vettoriali, ed A un operatore lineare di S in S'.

Ovviamente A è invertibile quando l'equazione nell'incognita $x \in S$:

$$(2) \qquad Ax = 0$$

ammette l'unica soluzione $x = 0$, cioè quando per $x \in S$ sussiste l'implicazione:

(1) Se non è soddisfatta tale condizione l'insieme (1) è vuoto. Invece quando la condizione è verificata risulta $D \neq \emptyset$ poiché almeno lo zero di S appartiene a D.

(2') $$Ax = 0 \Longrightarrow x = 0 \ .$$

Ciò significa che il nucleo di A è costituito dall'unico elemento 0, ed equivale al fatto che, qualunque sia $y \in A(S)$, l'equazione nell'incognita $x \in S$:

(3) $$Ax = y$$

ammette una e una sola soluzione (1).

Se A è invertibile, la corrispondenza:

$$y \in A(S) \longrightarrow \left(\text{l'unico } x \in S \text{ soluzione della (3)}\right)$$

è l'operatore inverso di A e si denota con A^{-1}; evidentemente A^{-1} è un operatore lineare [di S' in S e di dominio $A(S)$].

Per esprimere il fatto che A è invertibile, si dice anche che esiste l'inverso di A, oppure che esiste A^{-1}; in tale ipotesi, ovviamente esiste anche l'inverso di A^{-1}, ed è uguale all'operatore A.

Dalla definizione di A^{-1} consegue che:

$$Ax = y \iff x = A^{-1}y \qquad \forall x \in S \text{ e } \forall y \in A(S)$$

e che:

(4) $$A^{-1}Ax = x \quad \forall x \in S \ , \qquad AA^{-1}y = y \quad \forall y \in A(S) \ .$$

E' utile osservare che, qualunque sia l'operatore lineare $A : S \to S'$, si può presentare una (e una sola) delle seguenti eventualità :

I) $$\underline{\operatorname{Ker} A = \{0\} \quad \text{e} \quad A(S) = S'}$$

ossia

$$\underline{\text{esiste } A^{-1} \text{ ed è definito in tutto } S'}$$

ossia

$$\underline{\text{l'equazione } Ax = y \quad \text{ammette soluzione unica } \forall y \in S'} \ .$$

---

(1) Se la (3) ammettesse due soluzioni distinte x_1, x_2, l'elemento $x_1 - x_2$ di S sarebbe una soluzione della (2) diversa da 0.

II) $\ker A = \{0\}$ e $A(S) \neq S'$

ossia

esiste A^{-1} ma non e' definito in tutto S'

ossia

l'equazione $Ax = y$ ammette soluzione unica $\forall y \in A(S)$,

ma non ammette alcuna soluzione per almeno un $y \in S'$

III) $\ker A \neq \{0\}$

ossia

non esiste A^{-1}

ossia

l'equazione $Ax = y$ ammette infinite soluzioni $\forall y \in A(S)$ [1].

Si dimostra facilmente che:

(1.1) Se S, S' sono a dimensione finita ed hanno eguale dimensione, allora qualunque sia l'operatore lineare $A : S \to S'$ si ha che:

$$\ker A = \{0\} \iff A(S) = S'.$$

Pertanto nelle ipotesi poste su S ed S' la seconda delle predette eventualità non può verificarsi.

(1.2.) Qualunque siano gli operatori lineari A e B tali che abbia senso AB, si ha che:

Esistono A^{-1} e B^{-1} $\Longrightarrow$ Esiste $(AB)^{-1}$ e $(AB)^{-1} = B^{-1}A^{-1}$.

Osserviamo che la tesi del teor. (1.1) non sussiste se i due spazi non hanno la stessa dimensione. Ad esempio l'operatore lineare A che ad ogni $x \in \mathbf{R}$ associa la coppia $(x,x) \in \mathbf{R}^2$ è invertibile, ma il suo codominio non è tutto lo spazio $\mathbf{R}^2$ (il codominio è la bisettrice del primo e terzo quadrante).

Se un operatore lineare A : S→S' è invertibile ed ha per codominio tutto lo spazio, detti I ed I' rispettivamente l'operatore identico in S e in S' , l'inverso di A è l'unico operatore lineare $B : S' \to S$ tale che:

(5) $$BA = I \quad , \quad AB = I' \qquad ([2]).$$

Viceversa, *se esiste un operatore lineare B : S'→S che verifica le (5), allora A è invertibile ed ha per codominio tutto lo spazio, e il predetto B è l'inverso di A*: tale asserzione si deduce in base alle nozioni che saranno introdotte tra breve. Preliminarmente vogliamo evidenziare che una sola delle (5) non basta ad assicurare che A è invertibile ed ha per codominio tutto lo spazio, come risulta dal teorema che segue.

([1]) Se $Ker A \neq \{0\}$, esiste almeno un $x_0 \neq 0$ tale che $Ax_0 = 0$. Qualunque sia $y \in A(S)$, detto x^* un elemento di S tale che $Ax^* = y$, ogni x del tipo $x = x^* + tx_0$ con $t \in K$ è soluzione dell'equazione $Ax = y$; pertanto questa ha infinite soluzioni. Viceversa, se si verifica tale circostanza, è ovvio che $Ker A \neq \{0\}$.

([2]) Il fatto che per $B=A^{-1}$ sussistono le (5) consegue dalle (4) di pag.73. Quanto all'unicità, basta osservare che per la seconda delle (5) risulta $A(B(y))=y \ \forall \ y \in S'$, da cui segue $By=A^{-1}y \ \forall \ y \in S'$, ossia $B=A^{-1}$.

(1.3) Qualunque siano gli operatori lineari $L : S \to S'$, $L' : S' \to S$, sussiste l'implicazione:

$$L'L = I \Rightarrow \begin{cases} L \text{ è invertibile} \\ L' \text{ ha per codominio tutto lo spazio} \end{cases}$$

mentre dall'uguaglianza $L'L = I$ non consegue che L' è invertibile, e neppure che L ha per codominio tutto lo spazio

---

DIM. Essendo per ipotesi $x = L'(Lx)$ $\forall x \in S$, se x è un elemento di S e risulta $Lx = 0$ si ha $x = L'(Lx) = L'(0) = 0$. Dunque $Lx = 0 \Rightarrow x = 0$, ossia L è invertibile.

Fissato comunque $x \in S$, consideriamo l'equazione $L'y = x$ nell'incognita $y \in S'$. Posto $y = Lx$, si ha tenendo presente l'ipotesi: $L'y = L'Lx = x$; pertanto il punto Lx è soluzione della predetta equazione, e ciò prova che ogni punto dello spazio S appartiene al codominio di L'.

Per dimostrare la seconda parte del teorema consideriamo lo spazio $C^{(1)}([a,b])$ e denotiamo con L' l'operatore "derivazione", cioè l'operatore che ad ogni $y \in C^{(1)}([a,b])$ associa la derivata y', elemento dello spazio $C^{(0)}([a,b])$. Per ogni $x \in C^{(0)}([a,b])$ denotiamo poi con Lx la primitiva della funzione continua $x = x(t)$ nulla nel punto a: naturalmente Lx è un punto dello spazio $C^{(1)}([a,b])$.

Evidentemente l'operatore prodotto $L'L$ è l'identità in $C^{(0)}([a,b])$. Però L' non è invertibile (esistono infinite funzioni che hanno la stessa derivata), e l'operatore L ha per codominio la totalità delle funzioni $y \in C^{1}([a,b])$ nulle nel punto a, e perciò tale codominio non coincide con tutto lo spazio $C^{(1)}([a,b])$.

La (1.3) suggerisce le seguenti definizioni.

Sia A un operatore lineare di S in S'. Si dice che un operatore lineare:

$$B : S' \to S$$

è un <u>inverso sinistro</u> [risp. <u>destro</u>] di A, quando risulta:

$$BA = I \qquad [\text{risp. } AB = I'].$$

Dalla (1.3) consegue che:

(6) (<u>Esiste un inverso sinistro di A</u>) $\begin{cases} \Rightarrow \text{Esiste } A^{-1} \\ \not\Rightarrow A(S) = S' \end{cases}$

(7) (<u>Esiste un inverso destro di A</u>) $\begin{cases} \not\Rightarrow \text{Esiste } A^{-1} \\ \Rightarrow A(S) = S' \end{cases}$

In base alle precedenti implicazioni, verifichiamo che nel caso $S = S'$ si ha:

(<u>Esiste un unico inverso sinistro B di A</u>

[<u>risp. un unico inverso destro C di A</u>]) $\begin{cases} \Rightarrow \underline{\text{Esiste } A^{-1}} \\ \Rightarrow \underline{A(S) = S} \\ \Rightarrow \underline{B = A^{-1}} \ [\text{risp. } \underline{C = A^{-1}}] \end{cases}$.

Dim. Sia B l'unico inverso sinistro di A. Da $BA=I$ segue:

$$A(BA)=AI=A=IA,$$

ossia $(AB)A = IA$, cioé:

$$(8) \qquad (AB-I)A=0$$

avendo indicato con 0 l'operatore che ad ogni $x\in S$ fa corrispondere il vettore nullo di S. Dalla (8) segue (aggiungendo BA ad ambo i membri):

$$(AB-I+B)A=BA=I$$

sicché $AB-I+B$ é un inverso sinistro di A. Poiché per ipotesi l'unico inverso sinistro di A é B, ne segue $AB-I+B=B$, ossia $AB=I$. Pertanto B é anche un inverso destro di A. Da (6) e (7) consegue allora che esiste A^{-1} e che $A(S)=S$. Il fatto che $B=A^{-1}$ é ovvio. Analogamente si procede nell'ipotesi che esista un unico inverso destro C di A.

Osserviamo infine che:

(1.4) <u>Se esistono due inversi sinistri</u> [risp. <u>destri</u>] <u>di un operatore lineare $A: S \longrightarrow S'$, allora ne esistono infiniti</u>.

Dim. Siano B, C inversi sinistri di A, con $B\neq C$. Per ogni $\lambda\in K$ si ha allora:

$$[B+\lambda(C-B)]A=[B+\lambda C-\lambda B]A=BA+\lambda CA-\lambda BA=$$
$$=I+\lambda I-\lambda I=I$$

e quindi qulunque sia $\lambda\in K$ l'operatore $B+\lambda(C-B)$ é un inverso sinistro di A.

Analogamente si procede nell'ipotesi che esistano due inversi destri di A.

Capitolo IV - Spazi metrici . Spazi vettoriali topologici. Spazi vettoriali normati. Spazi pre-hilbertiani.

1. Spazi metrici.

Sia S un insieme . Un'applicazione d di $S \times S$ in $[0, +\infty[$ si chiama una metrica per S quando gode delle seguenti proprietà (assiomi della distanza):

1) $d(x,y) = 0 \iff x = y$ (assioma della coincidenza)

2) $d(x,y) = d(y,x) \quad \forall x, y \in S$ (assioma della simmetria)

3) $d(x,z) \leq d(x,y) + d(y,z) \quad \forall x, y, z \in S$ (assioma del triangolo)

Se d é una metrica per S ,la coppia (S,d) si chiama uno spazio metrico.

Sia (S,d) uno spazio metrico . Gli elementi di S si chiamano anche punti di S , e qualunque siano $x, y \in S$ il numero reale non negativo $d(x,y)$ si chiama distanza dei punti x e y .

Per ogni $x_0 \in S$ e per ogni $r \in [0, +\infty[$,l'insieme costituito dai punti $x \in S$ tali che $d(x, x_0) < r$ [risp. $d(x, x_0) \leq r$] si chiama cerchio aperto o sfera aperta [risp. cerchio chiuso o sfera chiusa] (di S), di centro x_0 e raggio r, e sarà da noi denotato con $C(x_0, r)$ (risp.con $C[x_0, r]$) .

(1.1.) Qualunque siano i cerchi aperti $C(x_1, r_1)$, $C(x_2, r_2)$ aventi intersezione non vuota , posto:

$$X = C(x_1, r_1) \cap C(x_2, r_2) ,$$

per ogni $x \in X$ esiste un cerchio aperto di centro x contenuto in X .

Dim. Per ogni $x \in X$, sia r_x il più piccolo dei numeri positivi

$$r_1 - d(x_1, x) \quad , \quad r_2 - d(x_2, x) .$$

E'immediato verificare che il cerchio $C(x, r_x)$ é contenuto in X .Infatti, per ogni $y \in C(x, r_x)$ si ha, per $i \in \{1, 2\}$:

$$d(x_i, y) \leq d(x_i, x) + d(x, y) < d(x_i, x) + [r_i - d(x_i, x)] = r_i ,$$

e quindi $y \in X$.

Sia $\mathcal{B}$ la classe dei cerchi aperti di (S,d). L'insieme S si può esprimere come unione di elementi di $\mathcal{B}$, in quanto S è l'unione dei cerchi aperti $C(x_0,n)$ con $n\in N$, essendo x_0 un fissato punto di S. Pertanto, per la (1.1) e per il fatto che $\varnothing\in\mathcal{B}$, sono soddisfatte le condizioni della (5.1) di pag.18 : ne segue che $\mathcal{B}$ è una base di topologia. La topologia per S di base $\mathcal{B}$, diciamola τ_d, si dice <u>generata</u> dalla metrica d, e lo spazio topologico (S,τ_d) si dice <u>generato</u> dallo spazio metrico (S,d). In (S,τ_d) un insieme è aperto quando è esprimibile come unione di cerchi aperti di (S,d).

Viceversa, assegnato uno spazio topologico (S,τ), non è detto che esso sia generato da uno spazio metrico, cioè che esista una metrica d per S tale che $\tau=\tau_d$; quando tale circostanza si verifica, si dice che lo spazio topologico (S,τ) è <u>metrizzabile</u> [1].

Se si conviene di identificare lo spazio metrico (S,d) con lo spazio topologico (S,τ_d) da esso generato, si può dire che gli spazi metrici sono particolari spazi topologici : precisamente quelli la cui topologia è generata da una metrica. Perciò con riferimento ad uno spazio metrico (S,d) si utilizzano tutte le nozioni introdotte nel Cap.I per gli spazi topologici, con l'intesa che esse si riferiscono allo spazio topologico (S,τ_d).

Evidentemente :

(1.2). <u>Se (S,d) è uno spazio metrico, per $x_0\in S$ e $I\subseteq S$ si ha :</u>

<u>I intorno di x_0</u> $\Longleftrightarrow$ <u>I contiene un cerchio aperto di centro x_0</u> .

Dimostriamo che :

(1.3). <u>Se (S,d) è uno spazio metrico, le proposizioni che seguono sono equivalenti :</u>

a) <u>(S,d) è separabile.</u>

b) <u>La topologia generata da d ha una base finita o numerabile.</u>

Dim. Poichè l'implicazione b$\Rightarrow$a è stata già dimostrata per una topologia qualunque [cfr. teor. (6.1) pag.19], si tratta di far vedere che a$\Rightarrow$b .

Se (S,d) è separabile, esiste un sottoinsieme Y di S, finito o numerabile, ovunque denso in S. Detto $\mathcal{B}$ l'insieme, che è finito o numerabile, costituito dalla parte vuota e dai cerchi aperti di (S,d) aventi il centro in un punto di Y e raggio del tipo $1/n$ con $n\in N$, proveremo che $\mathcal{B}$ è una base per la topologia generata da d.

Basta dimostrare che ogni cerchio aperto C_0 di (S,d) è unione di elementi di $\mathcal{B}$: si tratta di far vedere che ogni punto x di C_0 appartiene ad almeno un elemento B_x di $\mathcal{B}$ contenuto in C_0.

---

(1) Ad esempio, detto S un insieme qualunque, sia τ la topologia discreta per S, cioè la topologia costituita dall'insieme $\mathcal{P}(S)$ di tutte le parti di S. È immediato constatare che la funzione d definita in $S\times S$ ponendo :

$$d(x,y)=\begin{cases}0 & \text{se } x=y\\ 1 & \text{se } x\neq y\end{cases}$$

è una metrica per S, e che la topologia generata da d è proprio $\tau=\mathcal{P}(S)$. Pertanto lo spazio topologico (S,τ) è metrizzabile.

Detto $C(x,r)$ un cerchio aperto di centro x contenuto in C_0, e fissato $n \in N$ tale che $\frac{1}{n} < \frac{r}{2}$, esiste almeno un punto $y \in Y$ tale che $d(x,y) < \frac{1}{n}$. Posto allora $B_x = C(y, \frac{1}{n})$, ovviamente x appartiene a B_x, ed è facile constatare che $B_x \subset C(x,r)$, quindi che $B_x \subset C_0$: infatti per ogni $z \in B_x$ si ha:

$$d(z,x) \leq d(z,y) + d(y,x) < \frac{1}{n} + \frac{1}{n} = \frac{2}{n} < r,$$

sicché ogni punto di B_x appartiene a $C(x,r)$.

Se d e d' sono metriche per S, si dice che d e d' sono equivalenti quando esistono due numeri reali positivi h, K tali da risultare, qualunque siano $x, y \in S$:

(1) $$d'(x,y) \leq h\, d(x,y), \qquad d(x,y) \leq K\, d'(x,y).$$

In tale ipotesi, ogni cerchio aperto di (S,d) contiene un cerchio aperto di (S,d') e ogni cerchio aperto di (S,d') contiene un cerchio aperto di (S,d). Dopo ciò, ricordando che un insieme è aperto se e solo se è intorno di ogni suo punto, e tenendo presente la (1.2), è immediato che:

(1.4) Due metriche per S equivalenti generano la stessa topologia per S.

Evidentemente, ogni spazio metrico è uno spazio di Hausdorff. Evidente è altresì che in uno spazio metrico ogni cerchio chiuso è un insieme chiuso (in quanto il complementare, essendo intorno di ogni suo punto, è aperto).

Nello spazio R^n, la cui topologia è ovviamente generata dalla metrica usuale per R^n, la chiusura di un cerchio aperto è il cerchio chiuso avente lo stesso centro e lo stesso raggio, e l'interno di un cerchio chiuso è il cerchio aperto avente lo stesso centro e lo stesso raggio: avvertiamo che tale proposizione non è vera in tutti gli spazi metrici [*].

Sia (S,d) uno spazio metrico. Per ogni sottoinsieme X di S, si dice che X è limitato quando esiste un cerchio aperto di (S,d) contenente X. In tale ipotesi, se $X \neq \emptyset$ l'estremo superiore dell'insieme numerico descritto da $d(x',x'')$ al variare di (x',x'') in $X \times X$ si chiama diametro dell'insieme X. Qualunque siano i sottoinsiemi X, Y di S, non vuoti, l'estremo inferiore dell'insieme numerico descritto da $d(x,y)$ al variare di x in X e di y in Y, si chiama distanza di X e Y, e si denota con $d(X,Y)$. In particolare, per ogni $x \in S$ la distanza di $\{x\}$ e Y si denota con $d(x,Y)$.

Dimostriamo il seguente teorema:

[*] Ad es. non è vera nello spazio metrico che si ottiene associando ad un insieme S qualunque la metrica d considerata nella nota [*] di pag. 78.

(1.5) Sia (S,d) uno spazio metrico. Qualunque siano i sottoinsiemi C', C'' di S, chiusi e disgiunti, esistono due insiemi A', A'', aperti e disgiunti, tali che:

$$C' \subset A' \quad , \quad C'' \subset A'' \ .$$

Dim. Osserviamo in primo luogo che per ogni $x \in C'$ risulta $d(x,C'')>0$: infatti, se fosse $d(x,C'')=0$, ossia:

$$\inf_{c\in C''} d(x,c) = 0 \ ,$$

per ogni $\varepsilon>0$ esisterebbe $c \in C''$ tale che $d(x,c)<\varepsilon$; in altri termini, a ciascun intorno di x apparterrebbe almeno un punto di C'', ossia x apparterrebbe a $\overline{C''}=C''$ (il che è assurdo, in quanto $x\in C'$ e $C'\cap C''=\emptyset$). Analogamente, $d(y,C')>0 \ \forall y\in C''$.

Ciò premesso, per ogni $x \in C'$ sia A'_x il cerchio aperto di centro x e raggio $d(x,C'')/2$, e per ogni $y\in C''$ sia A''_y il cerchio aperto di centro y e raggio $d(y,C')/2$.

Posto:
$$A' = \bigcup_{x\in C'} A'_x \quad , \quad A'' = \bigcup_{y\in C''} A''_y$$

evidentemente A' è un aperto contenente C', e A'' un aperto contenente C''.

Proviamo che A' e A'' sono disgiunti. Procedendo per assurdo, ammettiamo che esista $z \in A'\cap A''$: allora esistono $x\in C'$ e $y\in C''$ tali che:

$$d(x,z)<\tfrac{1}{2}\,d(x,C'') \quad , \quad d(y,z)<\tfrac{1}{2}\,d(y,C') \ .$$

Se ad esempio $d(x,C'') \le d(y,C')$, si ha allora:

$$d(x,y) \le d(x,z)+d(z,y) < \tfrac{1}{2}\,d(x,C'')+\tfrac{1}{2}\,d(y,C') \le d(y,C') \ ,$$

ossia:

$$d(x,y) < d(y,C') = \inf_{c\in C'} d(c,y)$$

il che è assurdo.

(1.6) Siano (S,d) uno spazio metrico, e X un sottoinsieme di S.

Si ha che:

$$X \text{ compatto} \Longrightarrow X \text{ (chiuso e) limitato} \quad (1) \ .$$

---

(1) Si ricordi il Teor. (3.4) di pag. 14.

Dim. Scelto un punto x_0 di S , la famiglia di cerchi aperti :

$$(C(x_0,n))_{n\in N}$$

é ovviamente un ricoprimento aperto di S , e quindi anche di X . Essendo X compatto, tale famiglia contiene un ricoprimento finito di X , sia esso:

$$\{C(x_0,n_1),\dots,C(x_0,n_K)\}\ .$$

Posto $r=\max\{n_1,\dots,n_K\}$ si ha allora:

$$X\subseteq\bigcup_{i=1}^{K}C(x_0,n_i)=C(x_0,r)\ .$$

Essendo X contenuto in un cerchio di (S,d) , X é limitato.

Supponendo $X\subseteq R^n$, X é compatto se e solo se X é chiuso e limitato: avvertiamo che ciò non é vero in uno spazio metrico qualunque.

Consideriamo ora n spazi metrici:

$$(S_1,d_1)\ ,\ (S_2,d_2)\ ,\ \dots\ ,\ (S_n,d_n)$$

e denotiamo con S il prodotto cartesiano $S_1\times S_2\times\cdots\times S_n$.

Posto, qualunque siano $x=(x_1,\dots,x_n)\in S$, $y=(y_1,\dots,y_n)\in S$:

$$(2)\qquad d(x,y)=\sum_{i=1}^{n}d_i(x_i,y_i)$$

si definisce un'applicazione d di $S\times S$ in $[0,+\infty[$ che , come immediatamente si verifica, é una metrica per S . Lo spazio metrico (S,d) si chiama prodotto della famiglia di spazi metrici $((S_i,d_i))_{i\in\{1,\dots,n\}}$; é facile verificare che lo spazio topologico generato da (S,d) é il prodotto degli spazi topologici generati dalla famiglia suddetta.

Un semplice esempio di spazio metrico prodotto si ottiene considerando lo spazio metrico (R,d_0), essendo d_0 la metrica usuale per R definita ponendo $d_0(x,y)=|x-y|\ \ \forall(x,y)\in R\times R$. Il prodotto dello spazio metrico (R,d_0) per se stesso si ottiene associando all'insieme R^2 la metrica d definita ponendo:

$$d(x,y)=|x_1-y_1|+|x_2-y_2|\qquad \forall x=(x_1,x_2)\in R^2\ e\ \forall y=(y_1,y_2)\in R^2.$$

Notiamo che d non é la metrica usuale per R^2, diciamola d^*, che é definita ponendo:

$$d^*(x,y)=\sqrt{(x_1-y_1)^2+(x_2-y_2)^2}\qquad \forall x=(x_1,x_2)\in R^2\ e\ \forall y=(y_1,y_2)\in R^2.$$

Però le due metriche sono equivalenti. Infatti essendo $|x_1-y_1|=\sqrt{(x_1-y_1)^2}\le d^*(x,y)$ e analogamente $|x_2-y_2|\le d^*(x,y)$, risulta $d(x,y)\le 2\,d^*(x,y)$; d'altra parte dalla disuguaglianza [cfr. (2) pag. 54]:

$$(a+b)^\alpha\le 2^\alpha(a^\alpha+b^\alpha)\qquad (a\ge 0,\ b\ge 0,\ \alpha>0)$$

per $a=|x_1-y_1|^2$, $b=|x_2-y_2|^2$, $\alpha=1/2$ si deduce che $d^*(x,y)\le\sqrt{2}\,d(x,y)$.

Ovviamente lo stesso discorso può farsi considerando il prodotto di n spazi metrici uguali ad (R,d_0) : la metrica di tale spazio non coincide con la metrica usuale per R^n (ossia con la metrica euclidea), ma le due metriche sono equivalenti.

2. Spazi vettoriali topologici.

Detto S uno spazio vettoriale sul campo K, di sostegno Σ, consideriamo le due operazioni, di addizione e di moltiplicazione per uno scalare, associate a Σ:

$$\tau : (x,y) \in \Sigma \times \Sigma \to x+y \in \Sigma \quad , \quad \theta : (\lambda, x) \in K \times \Sigma \to \lambda x \in \Sigma .$$

Supponendo Σ munito di una topologia $\mathcal{T}$, e K munito della topologia usuale, sicché anche $\Sigma \times \Sigma$ e $K \times \Sigma$ si intendono come spazi topologici (1), non è detto che le operazioni τ e θ siano funzioni continue (2).

Se accade che τ e θ sono continue, si dice che la topologia $\mathcal{T}$ associata a Σ è compatibile con la struttura di S; in tale ipotesi, la coppia $(S, \mathcal{T})$ si chiama uno spazio vettoriale topologico.

Dunque per "spazio vettoriale topologico" si intende uno spazio topologico $(\Sigma, \mathcal{T})$, dove Σ è il sostegno di uno spazio vettoriale S e la topologia $\mathcal{T}$ è compatibile con la struttura di S.

Naturalmente tutte le nozioni introdotte nel Cap. I per gli spazi topologici si possono utilizzare anche con riferimento ad uno spazio vettoriale topologico.

Dimostriamo le proposizioni seguenti.

(2.1) Sia $(S, \mathcal{T})$ uno spazio vettoriale topologico, e siano:

$$I \subseteq S \qquad x, y \in S \qquad \lambda \in K - \{0\} .$$

Si ha che:

(1) $\quad I$ intorno di $x \iff I = x + I_0$, con I_0 intorno di 0

(2) $\quad I$ intorno di $x \implies \begin{cases} \lambda I \text{ intorno di } \lambda x \\ I + y \text{ intorno di } x+y \end{cases}$.

Dim. (1).

$\Rightarrow$ Sia $\bar{x} \in S$. Essendo $\mathcal{T}$ compatibile con la struttura di S, la funzione:

$$\varphi : x \in S \to x + \bar{x}$$

è continua (3): pertanto, se I è intorno di $\bar{x}$ $(= \varphi(0))$, esiste un intorno J di 0 tale che, per ogni $x \in S$:

$$(*) \qquad x \in J \Rightarrow \varphi(x) \in I .$$

In virtù della $(*)$ l'insieme $\ddot{\varphi}(I)$ contiene J intorno di 0, e quindi $\ddot{\varphi}(I)$ è un intorno di 0. Poiché:

---

(1) Si ricordi la nozione di "spazio topologico prodotto", introdotta a pag. 17.

(2) Ad es., se ad R si associa la topologia dell'es. 5 di pag. 29, non è difficile constatare che, considerando R come spazio vettoriale, le operazioni di addizione e di moltiplicazione in R non sono funzioni continue.

(3) In quanto composta mediante le funzioni continue:

$$f : x \in S \to (x, \bar{x}) \in S \times S \quad , \quad g : (x,y) \in S \times S \to x+y \in S .$$

$$x \in \check{\varphi}^{1}(I) \Leftrightarrow \varphi(x) \in I \Leftrightarrow x+\bar{x} \in I$$

posto $\check{\varphi}^{1}(I) = I_0$ ne segue $I = \bar{x} + I_0$.

$\Leftarrow$ Siano $\bar{x} \in S$ e I_0 un intorno di 0 .Poiché la funzione:

$$\psi : x \in S \longrightarrow x - \bar{x}$$

é continua, essendo I_0 un intorno di $0\,(=\psi(\bar{x}))$ esiste un intorno J di $\bar{x}$ tale che, per ogni $x \in S$:

$$x \in J \Rightarrow \psi(x) \in I_0 .$$

Conseguentemente l'insieme $\check{\psi}^{1}(I_0)$ contiene J e quindi é un intorno di $\bar{x}$.
Poiché $\check{\psi}^{1}(I_0) = \bar{x} + I_0$,si conclude che l'insieme $I = \bar{x} + I_0$ é un intorno di $\bar{x}$.

Dim. (2). Essendo $\mathcal{T}$ compatibile con la struttura di S , la funzione:

$$\varphi : x \in S \longrightarrow \lambda^{-1} x$$

é continua [1] ;pertanto, se I_0 é un intorno di $0\,(=\varphi(0))$, esiste un intorno J di 0 tale che,per ogni $x \in S$:

$$x \in J \Rightarrow \varphi(x) \in I_0 .$$

Conseguentemente $\check{\varphi}^{1}(I_0) \supseteq J$ intorno di 0 , e quindi $\check{\varphi}^{1}(I_0)$ é un intorno di 0 . Poiché :

$$x \in \check{\varphi}^{1}(I_0) \Leftrightarrow \varphi(x) \in I_0 \Leftrightarrow \lambda^{-1} x \in I_0$$

ne segue che $\check{\varphi}^{1}(I_0) = \lambda I_0$,e quindi λI_0 é un intorno di 0 .

Abbiamo così dimostrato la prima delle (2) nel caso $x = 0$ .Passando al caso generale, sia I un intorno di x , sicché $I = x + I_0$ con I_0 intorno di 0 (per la (1)) Si ha allora $\lambda I = \lambda x + \lambda I_0$; essendo λI_0 un intorno di 0 in virtù di quanto già acquisito, ancora per la (1) ne segue che λI é un intorno di λx .

Quanto alla seconda delle (2), basta osservare che,per la (1):

$$I \text{ intorno di } x \Rightarrow I = x + I_0 \text{ , con } I_0 \text{ intorno di } 0 \Rightarrow$$
$$\Rightarrow I + y = x + y + I_0 \Rightarrow I + y \text{ intorno di } x + y .$$

(2.2) Sia $(S, \mathcal{T})$ uno spazio vettoriale topologico, e siano $y \in S$, $\lambda \in K - \{0\}$
Sussistono le implicazioni:

X aperto $\Longrightarrow$ ($y + X$ aperto e λX aperto)
X chiuso $\Longrightarrow$ ($y + X$ chiuso e λX chiuso)
X compatto $\Longrightarrow$ ($y + X$ compatto e λX compatto)

Inoltre, qualunque sia $Y \subseteq S$:

$$(X \text{ aperto}) \Longrightarrow (X + Y \text{ aperto}) .$$

---

(1) In quanto composta mediante le funzioni continue:

$$f : x \in S \longrightarrow (\lambda^{-1}, x) \in K \times S , \quad g : (\mu, x) \in K \times S \longrightarrow \mu x \in S .$$

Dim. Nell'ipotesi X aperto, sia z un punto di $y+X$. Essendo $z-y \in X$, esiste un intorno I di $z-y$ contenuto in X. Posto $I=(z-y)+I_0$ con I_0 intorno di 0 (cfr.(1)), si ha dunque $(z-y)+I_0 \subseteq X$, e quindi $z+I_0 \subseteq y+X$. Pertanto $y+X$ è un intorno di z (in quanto $z+I_0$ è un intorno di z). Poichè $y+X$ è un intorno di ogni suo punto, $y+X$ è aperto.

Ancora nell'ipotesi X aperto, sia z un punto di λX, sicché $z=\lambda x$ con $x \in X$. Essendo X intorno di x, ne segue che λX è un intorno di λx (cfr.(2)), ossia un intorno di z. Poichè λX è intorno di ogni suo punto, λX è aperto.

Per dimostrare l'implicazione relativa al caso X chiuso, basta tener presente la precedente già acquisita, e osservare che:

$$\complement(y+X) = y+\complement X \qquad (1)$$

$$\complement \lambda X = \lambda \complement X$$

Quanto all'implicazione relativa al caso X compatto, basta osservare che, qualunque sia la famiglia $(A_i)_{i\in I}$ di parti di S, si ha:

$$(A_i)_{i\in I} \text{ ricoprimento di } y+X \implies (-y+A_i)_{i\in I} \text{ ricoprimento di } X$$
$$(A_i)_{i\in I} \text{ ricoprimento di } \lambda X \implies (\lambda^{-1}A_i)_{i\in I} \text{ ricoprimento di } X$$

e tener conto della prima implicazione dell'enunciato.

Per dimostrare l'ultima implicazione, basta osservare che:

$$X+Y = \bigcup_{y\in Y}(X+y)$$

e tener conto della prima implicazione.

(2.3) Sia $(S,\mathcal{T})$ uno spazio vettoriale topologico, e sia $X \subseteq S$. Si ha che:

$$X \text{ convesso} \implies (\overset{\circ}{X} \text{ e } \overline{X} \text{ convessi}).$$

Dim. Per provare che $\overset{\circ}{X}$ è convesso, si tratta di far vedere che per ogni $\lambda \in]0,1[$ il punto:

$$z = (1-\lambda)x + \lambda y$$

appartiene a $\overset{\circ}{X}$ se $x \in \overset{\circ}{X}$ e $y \in \overset{\circ}{X}$; tale asserzione, come vedremo, è vera anche se $x \in X$ e $y \in \overset{\circ}{X}$. In tali ipotesi, essendo y interno ad X esiste I_0 intorno di 0 tale che $y+I_0 \subseteq X$. Dimostriamo che:

$$z+\lambda I_0 \subseteq X.$$

---

(1) Per ogni sottoinsieme S_0 di S, con $\complement S_0$ si denota il complementare di S_0 rispetto ad S: non si può denotarlo con $-S_0$, in quanto S è uno spazio vettoriale e quindi $-S_0$ significa $(-1)S_0$.

Sia $u \in z + \lambda I_0$, e sia $i_0 \in I_0$ tale che $u = z + \lambda i_0$ .Si ha allora:

$$u = z + \lambda i_0 = (1-\lambda)x + \lambda y + \lambda i_0 = (1-\lambda)x + \lambda(y + i_0) .$$

Poichè x e $y + i_0$ sono elementi di X , e X è per ipotesi convesso, ne segue che $u \in X$.

In definitiva esiste un intorno di z , ossia $z + \lambda I_0$ (cfr.(1),e(2) con $x = 0$),che è contenuto in X :pertanto $z \in \overset{\circ}{X}$.

Per dimostrare che l'insieme $\overline{X}$ è convesso, siano x, y elementi di $\overline{X}$ e $\lambda \in]0,1[$ .Detto I un intorno del punto $[(1-\lambda)x + \lambda y]$, sia I_0 un intorno di 0 tale che (cfr. (1)):

$$I = (1-\lambda)x + \lambda y + I_0 .$$

Consideriamo un intorno J_0 di 0 tale che:

$$J_0 + J_0 \subseteq I_0 \quad (1) .$$

In virtù della (2), gli insiemi

$$(1-\lambda)^{-1} J_0 \quad , \quad \lambda^{-1} J_0$$

sono intorni di 0 ,e quindi, per la (1), gli insiemi:

$$x + (1-\lambda)^{-1} J_0 \quad , \quad y + \lambda^{-1} J_0$$

sono intorni rispettivamente di x e di y .

Poiché $x \in \overline{X}$ e $y \in \overline{X}$,esistono $x_1 \in X$, $y_1 \in X$ tali che:

$$x_1 \in x + (1-\lambda)^{-1} J_0 \quad , \quad y_1 \in y + \lambda^{-1} J_0 .$$

Si ha allora :

$$(1-\lambda)x_1 + \lambda y_1 \in [(1-\lambda)x + \lambda y] + J_0 + J_0 \subseteq$$
$$\subseteq [(1-\lambda)x + \lambda y] + I_0 = I .$$

Essendo $(1-\lambda)x_1 + \lambda y_1$ un punto di X (in quanto X è per ipotesi convesso), abbiamo dimostrato che a ciascun intorno I del punto $(1-\lambda)x + \lambda y$ appartiene almeno un punto di X , ossia che $(1-\lambda)x + \lambda y \in \overline{X}$.

Dopo ciò resta acquisito che $\overline{X}$ è un insieme convesso.

— — — — — — — — — —

(2.4) Sia (S, τ) uno spazio vettoriale topologico, e sia $S_0 \subseteq S$. Se S_0 è aperto, l'inviluppo convesso di S_0 è aperto.

Dim. Sia X l'inviluppo convesso di S_0 , e sia x un punto di X . Esistono allora $n \in N$, $x_1, \dots, x_n \in S_0$, $\lambda_1, \dots, \lambda_n$ numeri non negativi con $\lambda_1 + \dots + \lambda_n = 1$, tali che:

---

(1) Per provare l'esistenza di un siffatto intorno J_0 di 0 , basta invocare la continuità nel punto $(0,0)$ della funzione:

$$(x,y) \in S \times S \to x + y$$

(che è continua in quanto τ è compatibile con la struttura di S).

$$x = \lambda_1 x_1 + \cdots + \lambda_n x_n .$$

Denotati con $A_1, \ldots, A_n$ degli aperti contenuti in S_0 a cui appartengono rispettivamente $x_1, \ldots, x_n$, e posto:

$$A = \lambda_1 A_1 + \cdots + \lambda_n A_n ,$$

si ha $x \in A \subseteq X$. Essendo A un insieme aperto [per la (2.2)] contenuto in X a cui x appartiene, X e' un intorno di x. Dunque X e' un intorno di ogni suo punto, ossia e' aperto.

(2.5) Siano: X un sottoinsieme di uno spazio topologico, S uno spazio vettoriale topologico di Hausdorff, f e g funzioni definite in X a valori in S, λ un elemento di K.

Se f e g sono convergenti in un punto x_0 di accumulazione per X, si ha:

$$\lim_{x \to x_0} [f(x) + g(x)] = \lim_{x \to x_0} f(x) + \lim_{x \to x_0} g(x) \quad , \quad \lim_{x \to x_0} \lambda f(x) = \lambda \lim_{x \to x_0} f(x).$$

Di conseguenza, se f e g sono funzioni continue, anche $f+g$ e λf sono continue.

Si noti che senza l'ipotesi di Hausdorff la somma dei limiti va intesa come somma di sottoinsiemi di S,

Dim. La funzione $f+g$ e' composta mediante le funzioni:

$$x \in X \to (f(x), g(x)) \in S \times S \quad , \quad (u, v) \in S \times S \to u + v .$$

Detti l_1 ed l_2 i limiti di f e g in x_0, la prima funzione converge ovviamente in x_0 ad (l_1, l_2) [1], e la seconda funzione e' continua in quanto la topologia di S e' compatibile con l'operazione di addizione in S. Poiche' ovviamente sussiste il teorema sul limite di una funzione composta, ne consegue l'asserto relativo ad $f+g$.

Per la funzione λf basta applicare la definizione di limite, tenendo presente la (2) del teor. (2.1).

(2.6) Ogni spazio vettoriale topologico di Hausdorff e' connesso.

Dim. Procedendo per assurdo, ammettiamo che S sia uno spazio vettoriale topologico di Hausdorff non connesso, e denotiamo con $\{A, B\}$ una partizione di S costituita da due insiemi chiusi.

Fissati $x \in A$ e $y \in B$, consideriamo la funzione:

$$f : t \in [0,1] \to (1-t)x + ty \in S$$

Essendo f continua [2], i sottoinsiemi di $[0,1]$:

---

[1] Si ricordi la topologia dello spazio topologico prodotto [pag. 17].

[2] In quanto somma delle funzioni:

$$f_1 : t \in [0,1] \to x \quad , \quad f_2 : t \in [0,1] \to (-x)t \quad , \quad f_3 : t \in [0,1] \to yt ,$$

che sono ovviamente continue: la prima perche' e' costante, le altre due per la continuita' della moltiplicazione per uno scalare.

$$T' = \overset{-1}{f}(A), \quad T'' = \overset{-1}{f}(B)$$

sono chiusi in $[0,1]$, ossia sono chiusi. Inoltre dal fatto che $A \cup B = S$ [risp. $A \cap B = \phi$] consegue che $T' \cup T'' = [0,1]$ [risp. $T' \cap T'' = \phi$]. Essendo T' e T'' ovviamente non vuoti, si ha in definitiva che $\{T', T''\}$ è una partizione dell'intervallo $[0,1]$ costituita da insiemi chiusi: essendo ciò assurdo in quanto $[0,1]$ è connesso, l'asserto è provato.

— — — — — — — —

I sottospazi vettoriali di uno spazio vettoriale topologico non sono necessariamente chiusi. In proposito sussiste il seguente teorema, che ci limitiamo ad enunciare:

(2.7) Se $(S,\mathcal{T})$ è uno spazio vettoriale topologico di Hausdorff, ogni sottospazio vettoriale di S a dimensione finita è chiuso.

Anche quando un sottospazio vettoriale non è chiuso, la sua chiusura è a sua volta un sottospazio vettoriale. Infatti sussiste il teorema:

(2.8) Sia $(S,\mathcal{T})$ uno spazio vettoriale topologico, e sia $H \subseteq S$. Si ha che:
H sottospazio vettoriale di $S \Longrightarrow \overline{H}$ sottospazio vettoriale di S.

Dim. Siano x, y punti di $\overline{H}$, e sia I un intorno di $x+y$. Detto I_0 un intorno di 0 tale che:

$$I = (x+y) + I_0$$

consideriamo un intorno J_0 di 0 tale che:

$$J_0 + J_0 \subseteq I_0 \quad (1).$$

Poiché $x + J_0$ e $y + J_0$ sono intorni rispettivamente di x e di y, esistono $x_1 \in H$ e $y_1 \in H$ tali che:

$$x_1 \in x + J_0 \quad , \quad y_1 \in y + J_0$$

e quindi tali che:

$$x_1 + y_1 \in (x+y) + J_0 + J_0 \subseteq (x+y) + I_0 = I .$$

Essendo $x_1 + y_1$ un punto di H, abbiamo visto che a ciascun intorno I di $(x+y)$ appartiene almeno un punto di H, ossia che $x+y \in \overline{H}$.

Resta da far vedere che per ogni $x \in \overline{H}$ e per ogni $\lambda \in K$ risulta $\lambda x \in \overline{H}$. Essendo ciò evidente se $\lambda = 0$, supponiamo $\lambda \neq 0$.

Detto I un intorno di λx, poniamo $I = \lambda x + I_0$, con I_0 intorno di 0. Essendo $\lambda^{-1} I_0$ un intorno di 0 [per la (2.1)], ne segue che $x + \lambda^{-1} I_0$ è un intorno di x, e quindi esiste $x_1 \in H$ tale che $x_1 \in x + \lambda^{-1} I_0$, ossia tale che $\lambda x_1 \in \lambda x + I_0 = I$. Poiché λx_1 è un punto di H, si conclude che a ciascun intorno I di λx appartiene almeno un punto di H, ossia che $\lambda x \in \overline{H}$.

(1) Cfr. nota (1) a piè di pag. 85.

Lo spazio vettoriale topologico $(S,\mathcal{T})$ si dice localmente convesso quando ogni intorno di 0 contiene un intorno convesso di 0. Ciò equivale al fatto che, qualunque sia $x \in S$, ogni intorno di x contiene un intorno convesso di x: all'uopo basta ricordare che ogni intorno I di x é del tipo $I = x + I_0$ con I_0 intorno di 0 [cfr. (2.1)], e osservare che, qualunque sia $Y \subseteq S$:

$$Y \text{ convesso} \implies x+Y \text{ convesso.}$$

Terminiamo questo numero con la seguente definizione.

Se $(S,\mathcal{T})$ é uno spazio vettoriale topologico, e H é un sottospazio vettoriale di S, la topologia indotta da $\mathcal{T}$ su H, diciamola $\mathcal{T}_H$, é compatibile con la struttura di H; lo spazio $(H,\mathcal{T}_H)$ si chiama sottospazio vettoriale topologico di $(S,\mathcal{T})$, e si denota semplicemente con H.

3. Spazi vettoriali normati.

Sia S uno spazio vettoriale, e sia ν un'applicazione di S in $[0,+\infty[$. Si dice che ν é una norma per S quando, posto:

$$\nu(x) = \|x\| \qquad \forall x \in S$$

sono verificate le seguenti condizioni:

a) $$\|x\| = 0 \iff x = 0$$

b) $$\|\lambda x\| = |\lambda| \cdot \|x\| \qquad \forall \lambda \in K \text{ e } \forall x \in S$$

c) $$\|x+y\| \leq \|x\| + \|y\| \qquad \forall x,y \in S.$$

La c) si chiama proprietà triangolare della norma.

Quando occorre precisare che la norma ν é definita nello spazio vettoriale S, per ogni $x \in S$ si pone:

$$\nu(x) = \|x\|_S .$$

Se ν é una norma per S, la coppia (S,ν) si chiama uno spazio vettoriale normato.

Sia (S,ν) uno spazio vettoriale normato. Ponendo, qualunque siano $x,y \in S$:

(1) $$d(x,y) = \|x-y\|_S$$

si definisce un'applicazione d di $S \times S$ in $[0,+\infty[$ che, come immediatamente si verifica, é una metrica per S: conseguentemente essa genera una topologia per S, che dicesi generata dalla norma ν.

Dimostriamo che:

(3.1) La topologia $\mathcal{T}_\nu$ per S generata dalla norma ν è compatibile con la struttura di S.

Dim. Si tratta di far vedere che le funzioni :

$$f:(x,y)\in S\times S\to x+y \quad , \qquad g:(\lambda,x)\in K\times S\to \lambda x$$

sono continue.

Sia $(x_0,y_0)\in S\times S$,e sia I un intorno di $f(x_0,y_0)=x_0+y_0$. Esiste allora [cfr. proposiz.(1.2)] un cerchio aperto di (S,d) di centro x_0+y_0 ,diciamolo $C(x_0+y_0,\varepsilon)$, contenuto in I . L'insieme J costituito dai punti $(x,y)\in S\times S$ tali che:

$$\|x-x_0\|+\|y-y_0\|<\varepsilon$$

è un intorno di (x_0,y_0) (1), e per ogni $(x,y)\in J$ risulta:

$$\|f(x,y)-f(x_0,y_0)\|=\|(x+y)-(x_0+y_0)\|\leq\|x-x_0\|+\|y-y_0\|<\varepsilon ,$$

ossia $f(x,y)\in C(x_0+y_0,\varepsilon)\subseteq I$ .Pertanto f è continua nel punto (x_0,y_0).

Consideriamo ora un punto $(\lambda_0,x_0)\in K\times S$,e denotiamo con I un intorno di $g(\lambda_0,x_0)=\lambda_0 x_0$. Esiste allora un cerchio aperto di (S,d) di centro $\lambda_0 x_0$,diciamolo $C(\lambda_0 x_0,\varepsilon)$,contenuto in I .L'insieme J costituito dai punti $(\lambda,x)\in K\times S$ tali che:

$$|\lambda-\lambda_0|+\|x-x_0\|<\frac{\varepsilon}{1+\varepsilon+|\lambda_0|+\|x_0\|}$$

è un intorno di (λ_0,x_0) ,e per ogni $(\lambda,x)\in J$ risulta:

$$\|g(\lambda,x)-g(\lambda_0,x_0)\|=\|\lambda x-\lambda_0 x_0\|=\|\lambda(x-x_0)+(\lambda-\lambda_0)x_0\|\leq$$

$$\leq|\lambda|\cdot\|x-x_0\|+|\lambda-\lambda_0|\cdot\|x_0\|<(1+\varepsilon+|\lambda_0|+\|x_0\|)(\|x-x_0\|+|\lambda-\lambda_0|)<\varepsilon ,$$

ossia $g(\lambda,x)\in C(\lambda_0 x_0,\varepsilon)\subseteq I$ .Pertanto g è continua nel punto (λ_0,x_0)

In virtù della (3.1) ,lo spazio $(S,\mathcal{T}_\nu)$ è uno spazio vettoriale topologico, che

---

(1) Si tenga presente quanto detto a proposito della (2) del n.1 (pag.81).

dicesi generato da V. Se si conviene di identificare lo spazio (S, τ_ν) con lo spazio metrico (S, d) [dove d è la metrica definita mediante la (1)], ed entrambi con lo spazio vettoriale normato (S, ν), si può dire che gli spazi vettoriali normati sono particolari spazi vettoriali topologici [precisamente quelli la cui topologia è generata da una norma], nonché particolari spazi metrici [precisamente quelli la cui metrica è definita mediante una norma].

Con riferimento allo spazio vettoriale normato (S, ν) si possono utilizzare tutte le nozioni introdotte per gli spazi topologici, per gli spazi metrici e per gli spazi vettoriali topologici, nonché i relativi risultati, con l'intesa che si fa riferimento rispettivamente allo spazio topologico (Σ, τ_ν), dove Σ è il sostegno di S, allo spazio metrico (Σ, d), dove d è la metrica definita mediante la (1), e allo spazio vettoriale topologico (S, τ_ν). In particolare, possiamo dire che (S, ν) è uno spazio di Hausdorff, ed è uno spazio vettoriale topologico localmente convesso, in quanto ogni intorno di 0 contiene un cerchio aperto di centro 0, e d'altra parte ogni cerchio di (S, ν) è un insieme convesso [1].

D'ora in avanti, quando non vi è possibilità di equivoci, in luogo di (S, ν) scriveremo semplicemente S.

Dimostriamo la proposizione seguente che, come abbiamo già avvertito, non è vera in tutti gli spazi metrici:

(3.2) *In ogni spazio vettoriale normato S, la chiusura di un cerchio aperto [l'interno di un cerchio chiuso] è il cerchio chiuso [aperto] che ha lo stesso centro e lo stesso raggio, supposto non nullo.*

In simboli, qualunque siano $x_0 \in S$ ed $r > 0$, si ha:

$$\overline{C(x_0, r)} = C[x_0, r] \tag{2}$$

$$\big(C[x_0, r]\big)^\circ = C(x_0, r). \tag{3}$$

Dim. Poiché in un qualunque spazio metrico si ha $\overline{C(x_0, r)} \subseteq C[x_0, r]$, in quanto

(1) Invero, se x', x'' appartengono al cerchio aperto $C(x_0, r)$, per ogni $\lambda \in [0,1]$ si ha:
$\| (1-\lambda)x' + \lambda x'' - x_0 \| = \| (1-\lambda)(x'-x_0) + \lambda(x''-x_0) \| \leq (1-\lambda)\|x'-x_0\| + \lambda\|x''-x_0\| < (1-\lambda)r + \lambda r = r$
e quindi $(1-\lambda)x' + \lambda x'' \in C(x_0, r)$. Analogamente si procede per un cerchio chiuso.

$C[x_0,r]$ è un insieme chiuso contenente $C(x_0,r)$, per dimostrare la (2) si tratta di far vedere che sussiste l'inclusione opposta, ossia che ogni punto $x \in C[x_0,r]$ è aderente a $C(x_0,r)$.

Detto I un intorno di x , sia $C(x,\delta)$ un cerchio aperto di centro x contenuto in I , con $\delta<r$ .Scelto $\lambda \in]1-\frac{\delta}{r}, 1[$, consideriamo il punto di S :

$$y=(1-\lambda)x_0+\lambda x .$$

Risultando

$$\|y-x\| = (1-\lambda)\|x_0-x\| \leq (1-\lambda)r < \delta$$
$$\|y-x_0\| = \lambda\|x-x_0\| \leq \lambda r < r ,$$

y è un punto di I (in quanto $y \in C(x,\delta) \subseteq I$) e appartiene a $C(x_0,r)$: pertanto x è aderente a $C(x_0,r)$, e con ciò la (2) è dimostrata.

Per provare la (3), essendo $C(x_0,r) \subseteq (C[x_0,r])^\circ$, si tratta di far vedere che sussiste l'inclusione opposta, ossia che ogni punto interno a $C[x_0,r]$ appartiene a $C(x_0,r)$: procedendo per assurdo, ammettiamo che esista un punto x interno a $C[x_0,r]$ non appartenente a $C(x_0,r)$, sicché $\|x-x_0\|=r$.

Essendo x interno a $C[x_0,r]$, esiste $\rho>0$ tale che:

(4) $$C(x,\rho) \subset C[x_0,r] .$$

Scelto un numero positivo $\lambda < \rho/r$,consideriamo il punto di S :

$$y = x+\lambda(x-x_0) .$$

Risultando:

$$\|y-x\| = \lambda\|x-x_0\| = \lambda r < \rho$$
$$\|y-x_0\| = (1+\lambda)\|x-x_0\| = (1+\lambda)r > r$$

il punto y appartiene a $C(x,\rho)$ e non appartiene a $C[x_0,r]$,il che è in contrasto con la (4).

Osserviamo che :

(3.3) Per ogni spazio vettoriale normato S , qualunque siano $x_0 \in S$ e $r>0$ si ha:

$$C(x_0,r) = x_0 + C(0,r) \quad , \quad C[x_0,r] = x_0 + C[0,r] .$$

Dim. Qualunque sia $x \in S$, si ha infatti:

$$x \in x_0+C(0,r) \Leftrightarrow x-x_0 \in C(0,r) \Leftrightarrow \|x-x_0\|<r$$
$$x \in x_0+C[0,r] \Leftrightarrow \|x-x_0\| \leq r .$$

(3.4) Per ogni spazio vettoriale normato S, e per ogni $X \subseteq S$, le seguenti proposizioni sono equivalenti:

1) X è limitato

2) Esiste $k>0$ tale che:

(5) $\quad \|x\| \leq k \quad \forall x \in X$.

Dim. Se X è limitato, esiste un cerchio aperto $C(x_0, r)$ contenente X. Per ogni $x \in X$ si ha allora $\|x - x_0\| < r$, e quindi $\|x\| \leq \|x - x_0\| + \|x_0\| < r + \|x_0\|$

Reciprocamente, se vale la (5), ciò significa che X è contenuto nel cerchio $C[0,k]$, e quindi X è limitato.

Sia S uno spazio vettoriale, e siano ν', ν'' norme per S. Si dice che ν' e ν'' sono norme equivalenti quando esistono due numeri reali positivi h, k tali da risultare, qualunque sia $x \in S$:

$$\nu'(x) \leq h\,\nu''(x) \quad , \quad \nu''(x) \leq k\,\nu'(x) .$$

E' immediato constatare che:

(3.5) Due norme equivalenti per S generano la stessa topologia per S.

Dim. Basta osservare che nell'ipotesi posta le metriche per S:

$$d : (x,y) \in S \times S \to \nu(x-y) \quad , \quad d' : (x,y) \in S \times S \to \nu'(x-y)$$

sono equivalenti, e ricordare la proposizione (1.4) a pag 79.

Sia (S,ν) uno spazio vettoriale normato, e sia H un sottospazio vettoriale di S. Ovviamente, la restrizione ν_H di ν ad H é una norma per H, e la topologia per H da essa generata é quella indotta dalla topologia di S su H. Lo spazio vettoriale normato (H,ν_H) dicesi sottospazio (vettoriale normato) di (S,ν) di sostegno H.

Detta τ la topologia per S generata da ν, e τ_H la topologia per H generata da ν_H, lo spazio (H,τ_H) é il sottospazio vettoriale topologico di (S,τ) di sostegno H.

Consideriamo ora una famiglia finita $(S_i)_{i\in\{1,..,k\}}$ di spazi vettoriali normati, e denotiamo con S lo spazio vettoriale prodotto $S_1\times S_2\times\cdots\times S_k$.

Ponendo, per ogni $x=(x_1,..,x_k)\in S$:

$$\nu(x)=\sum_{i=1}^{k}\|x_i\|_{S_i}$$

si definisce un'applicazione di S in $[0,+\infty[$ che, come facilmente si verifica, é una norma per S. Lo spazio vettoriale normato (S,ν) si chiama prodotto della famiglia di spazi vettoriali normati sopra considerata. Detta d la metrica per S definita mediante ν, e d_i $(i=1,..,k)$ la metrica per S_i definita mediante la norma per S_i, ovviamente (S,d) é il prodotto degli spazi metrici della famiglia $(S_i,d_i)_{i\in\{1,..,k\}}$.

Sia (S,ν) uno spazio vettoriale normato, e sia H un sottospazio vettoriale di S. Si dimostra che, se H é chiuso, l'applicazione:

$$\nu(H):\ X\in S/H\longrightarrow \inf_{x\in X}\|x\|_S$$

é una norma per S/H. Lo spazio vettoriale normato $(S/H,\nu(H))$ si chiama spazio quoziente di (S,ν) rispetto ad H. Quando non vi é possibilità di equivoci tale spazio si denota semplicemente con S/H.

4. Spazi pre-hilbertiani

Sia S uno spazio vettoriale sul campo reale [risp. sul campo complesso], e sia $p(x,y)$ una forma bilineare [risp. sesquilineare] in S.

Si dice che p é un prodotto scalare per S, quando:

a) $p(x,y)$ é simmetrica.

b) $p(x,x)>0\quad\forall x\in S-\{0\}$.

Se p é un prodotto scalare per S, la coppia (S,p) si chiama uno spazio pre-hilbertiano (reale [risp.complesso] se é uno spazio vettoriale sul campo reale [risp. complesso]), di sostegno S.

Quando si considera uno spazio pre-hilbertiano (S,p), esso si denota semplicemente con S (purché non vi sia possibilità di equivoci), e qualunque siano $x,y \in S$ il numero $p(x,y)$ si denota col simbolo $(x,y)_S$, o semplicemente con

$$(x,y)$$

se non vi é possibilità di equivoci.

Il numero $(x,y)_S$ si chiama prodotto scalare di x e y (in S)
Sussiste la seguente disuguaglianza (disuguaglianza di Schwarz), qualunque siano $x,y \in S$:

$$(1) \qquad |(x,y)| \leqslant \sqrt{(x,x)}\ \sqrt{(y,y)} \ .$$

Dim. Fissati $x,y \in S$, possiamo supporre $x \neq 0$, $y \neq 0$ e $(x,y) \neq 0$, in quanto la (1) é banale se $x=0$, o $y=0$, oppure $(x,y)=0$.
In virtù della condizione b), qualunque siano $t \in R$ e $\lambda \in K$ risulta:

$$(tx+\lambda y, tx+\lambda y) \geqslant 0$$

ossia:

$$(x,x)t^2+[(x,y)\overline{\lambda}+\overline{(x,y)}\lambda]t+(y,y)|\lambda|^2 \geqslant 0 \ .$$

Pertanto, posto

$$\lambda = \frac{(x,y)}{|(x,y)|}$$

qualunque sia $t \in R$ risulta:

$$(*) \qquad (x,x)t^2+2|(x,y)|t+(y,y) \geqslant 0 \ .$$

Poiché il primo membro della (*) é un polinomio di secondo grado in t, la (*) stessa implica che il discriminante di tale polinomio é non positivo.

Si ha dunque :

$$|(x,y)|^2-(x,x)(y,y) \leqslant 0$$

ossia la (1).

(4.1) Qualunque sia lo spazio pre-hilbertiano (S, p), l'applicazione:

$$(2) \qquad \nu : x \in S \to \sqrt{(x,x)}$$

è una norma per S.

Dim. Si tratta di dimostrare che ν gode della proprietà triangolare, in quanto le rimanenti condizioni affinché ν sia un norma sono ovvie:

Qualunque siano $x, y \in S$ essendo:

$$\nu^2(x+y) = (x+y, x+y) = (x,x) + (x,y) + \overline{(x,y)} + (y,y) \leq$$
$$\leq (x,x) + 2\,|(x,y)| + (y,y)$$

in virtù della disuguaglianza di Schwarz risulta:

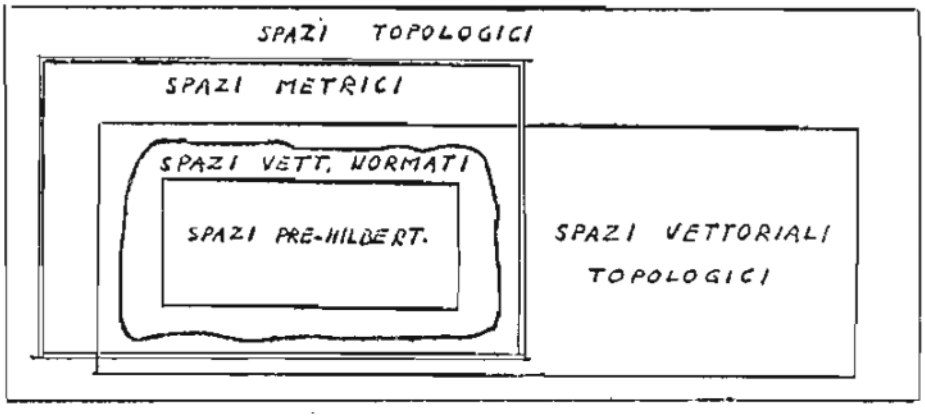
$$\nu^2(x+y) \leq \nu^2(x) + 2\,\nu(x)\,\nu(y) + \nu^2(y) = \big(\nu(x)+\nu(y)\big)^2$$

e quindi:

$$\nu(x+y) \leq \nu(x) + \nu(y) .$$

Se (S,p) è uno spazio pre-hilbertiano, la norma ν per S definita dalla (2) si dice generata dal prodotto scalare p, e lo spazio vettoriale normato (S, ν) si dice generato da p. Se si conviene di identificare lo spazio (S,p) con lo spazio vettoriale normato (S,ν), si può dire che gli spazi pre-hilbertiani sono particolari spazi vettoriali normati: precisamente quelli la cui norma è generata da un prodotto scalare.

Adottando la convenzione suddetta, nonché le convenzioni di cui si è parlato nei n. 1, 2, 3, riassumiamo gli spazi introdotti con lo schema seguente.



Con riferimento allo spazio pre-hilbertiano (S,p) si possono utilizzare tutte le nozioni e i risultati relativi agli spazi topologici, metrici, vettoriali topologici e vettoriali normati, intendendo che si fa riferimento rispettivamente agli spazi (Σ,τ), (Σ,d), (S,τ), (S,ν), dove Σ è il sostegno di S, ν la norma (2), d la metrica definita mediante ν, e τ la topologia generata da d.

Considerato lo spazio vettoriale K^n, osserviamo che, se si pone per ogni $x=(x_1,\dots,x_n)\in K^n$ e per ogni $y=(y_1,\dots,y_n)\in K^n$:

$$p(x,y)=\sum_{i=1}^{n} x_i\,\overline{y_i}$$

si definisce una forma sesquilineare in K^n (bilineare se $K=R$) che, come facilmente si verifica, è un prodotto scalare per K^n. In particolare, nello spazio vettoriale R^3, $p(x,y)$ è il prodotto scalare dei vettori x,y rispettivamente di componenti (x_1,x_2,x_3) e (y_1,y_2,y_3): pertanto se $x\neq 0$ e $y\neq 0$ il numero reale

$$\frac{p(x,y)}{\|x\|\cdot\|y\|} \quad (1)$$

rappresenta il coseno dell'angolo compreso tra 0 e π formato dai vettori x e y; tali vettori sono ortogonali se e solo se $p(x,y)=0$.

Quanto sopra suggerisce le seguenti definizioni.

Sia S uno spazio pre-hilbertiano. Qualunque siano $x,y\in S$, si dice che x e y sono <u>ortogonali</u> quando risulta:

$$(x,y)=0 \ .$$

Per ogni $Y\subseteq S$, $Y\neq\emptyset$ si dice che l'elemento x di S è <u>ortogonale ad Y</u> quando x è ortogonale a ciascun elemento di Y ; l'insieme costituito dagli elementi di S ortogonali ad Y si chiama il <u>complemento ortogonale</u> di Y , e si denota con

$$Y^{\perp} \ .$$

Osserviamo che, se x,y sono elementi di S <u>ortogonali</u>, risulta:

$$(3)\qquad \|x+y\|^2=\|x\|^2+\|y\|^2 \quad (2) \ .$$

Nel caso $S=\mathbf{R}^2$ (munito della norma usuale) la (3) esprime il teorema di Pitagora. Notiamo che se lo spazio normato $\mathbf{R}^2$ si considera come prodotto dello spazio normato $\mathbf{R}$ per se stesso, la norma (equivalente a quella usuale) non è generata da alcun prodotto scalare: basta constatare che, essendo $\|x\|=|x_1|+|x_2|\ \ \forall (x_1,x_2)\in\mathbf{R}\times\mathbf{R}$, per i vettori ortogonali $x=(1,1)$ e $y=(1,-1)$ non è verificata la (3).

---

(1) Abbiamo posto $\|x\|=\nu(x)$ con ν data dalla (2), ossia $\|x\|=\sqrt{p(x,x)}$. Ovviamente nel caso in esame $\|x\|$ rappresenta il modulo del vettore di componenti x_1,x_2,x_3. Inoltre il numero reale $p(x,y)/\|y\|$ rappresenta la componente di x secondo la direzione di y.

(2) Infatti $\|x+y\|^2=(x+y,x+y)=(x,x)+(x,y)+(y,x)+(y,y)=\|x\|^2+\|y\|^2$.

Osserviamo altresì che :

$$(4) \qquad \|x+y\|^2+\|x-y\|^2 = 2\left(\|x\|^2+\|y\|^2\right) \quad \forall x,y \in S \quad (1).$$

Nel caso $S=R^2$, la (4) esprime il fatto che in un parallelogramma la somma dei quadrati delle lunghezze delle diagonali é uguale alla somma dei quadrati delle lunghezze dei lati.

Si dimostra che, se (S,ν) è un qualunque spazio vettoriale normato, con S spazio vettoriale sul campo reale , e se la norma ν in S é tale da verificare la (4), allora posto:

$$p(x,y) = \frac{1}{4}\left(\|x+y\|^2-\|x-y\|^2\right) \qquad \forall x,y \in S$$

si definisce un'applicazione di $S\times S$ in R che é un prodotto scalare per S , e la norma per S generata da p coincide con ν . In altri termini, nell'ipotesi posta lo spazio vettoriale normato (S,ν) é necessariamente generato da uno spazio pre-hilbertiano.

Pur essendo banale, rileviamo che la diseguaglianza di Schwarz (cfr.(1)) si esprime come segue :

$$(5) \qquad |(x,y)| \leq \|x\|\cdot\|y\| \qquad \forall x,y \in S .$$

Osserviamo infine che:

(4.2) Qualunque sia lo spazio pre-hilbertiano S , per ciascun $y \in S$ il funzionale (complesso se S é complesso):

$$F_y : x \in S \to (x,y)$$

é continuo (2).

Dim. L'asserto é banale se $y=0$. Supposto $y \neq 0$,qualunque sia $x_0 \in S$ si ha, utilizzando la diseguaglianza di Schwarz :

$$|F_y(x)-F_y(x_0)| = |F_y(x-x_0)| = |(x-x_0,y)| \leq \|x-x_0\|\cdot\|y\| .$$

Dopo ciò, per ogni $\varepsilon > 0$, detto I il cerchio aperto di S di centro x_0 e raggio $\varepsilon/\|y\|$, si ha che:

$$x \in I \Longrightarrow |F_y(x)-F_y(x_0)| < \varepsilon .$$

---

(1) Infatti $\|x+y\|^2+\|x-y\|^2 = (x+y,x+y)+(x-y,x-y) = 2(x,x)+2(y,y) = 2\|x\|^2+2\|y\|^2$.

(2) Nel seguito il risultato espresso dalla (4.2) potrà essere dedotto direttamente dalla (5), utilizzando la (2.1) di pag. 138.

5. Alcuni spazi funzionali. Spazi di successioni.

Quasi tutti gli spazi vettoriali considerati nelle pagine 54-61 solitamente si intendono, salvo esplicito avviso contrario, muniti di un'opportuna topologia (topologia usuale) che spesso deriva da una norma (norma usuale); talvolta poi la norma usuale è generata da un prodotto scalare.

In questo numero elenchiamo tali spazi, in un ordine opportuno anche se diverso da quello seguito nelle pagine sopra citate, riportando la norma usuale quando sono normati, e qualche informazione sulla topologia negli altri casi.

Eventuali affermazioni relative agli spazi saranno riportate senza dimostrazione.

Preliminarmente ricordiamo che allo spazio vettoriale K^n è associato il prodotto scalare definito da:

$$(x,y)_{K^n} = \sum_{i=1}^{n} x_i \overline{y_i} \qquad \left(x=(x_1,\dots,x_n),\ y=(y_1,\dots,y_n) \right)$$

sicché la norma usuale in K^n è:

$$\|x\|_{K^n} = \left(\sum_{i=1}^{n} |x_i|^2 \right)^{1/2} \qquad \left(x=(x_1,\dots,x_n) \right)$$

e genera la metrica euclidea, definita da:

$$d(x,y) = \left(\sum_{i=1}^{n} |x_i - y_i|^2 \right)^{1/2} \qquad \left(x=(x_1,\dots,x_n),\ y=(y_1,\dots,y_n) \right).$$

In particolare la norma usuale nello spazio K è il modulo (ossia: norma di x = modulo di x).

Se si considera K^n come prodotto di n spazi normati uguali a K, la norma in K^n è:

$$\|x\| = \sum_{i=1}^{n} |x_i| \qquad \left(x=(x_1,\dots,x_n) \right)$$

ed è equivalente alla norma usuale in K^n, in quanto risulta (denotando con $\|\cdot\|$ la norma usuale e con $\|\cdot\|_1$ l'altra):

$$\|x\|_1 = \sum_{i=1}^{n} \sqrt{|x_i|^2} \le \sum_{i=1}^{n} \|x\| = n\|x\|$$

nonché (denotando con e_i il vettore di K^n avente la componente i^{ma} uguale ad 1 e le altre uguali a 0):

$$\|x\| = \|x_1 e_1 + \dots + x_n e_n\| \le |x_1| \cdot \|e_1\| + \dots |x_n| \cdot \|e_n\| = \|x\|_1 .$$

Vogliamo qui precisare che, quando consideriamo lo spazio vettoriale K^n, intendiamo lo spazio vettoriale C^n sul campo C oppure lo spazio vettoriale R^n sul campo R. In tal senso lo spazio vettoriale C^n ha dimensione n, come R^n. Notiamo che se si considera lo spazio vettoriale C^n sul campo R, la dimensione è $2n$ (1).

---

(1) Ogni vettore dello spazio è combinazione lineare dei $2n$ vettori:

$$e_1 = (1,0,0,\dots,0),\ e_2 = (0,1,0,0,\dots,0),\ \dots,\ e_n = (0,0,0,\dots,0,1)$$
$$\varepsilon_1 = (j,0,0,\dots,0),\ \varepsilon_2 = (0,j,0,0,\dots,0),\ \dots,\ \varepsilon_n = (0,0,0,\dots,0,j) \qquad (j=\text{unità immaginaria})$$

ed è facile constatare che tali vettori sono linearmente indipendenti.

Elenchiamo ora gli spazi accennati all'inizio.

1) ℓ_∞ ed $L^\infty(\Omega)$ sono spazi vettoriali normati, con le norme definite da:

$$\|x\|_{\ell_\infty} = \sup_{k\in N} |x_k| \quad [x=(x_k)_{k\in N}], \qquad \|x\|_{L^\infty(\Omega)} = \sup_{\Omega} |x| \quad \text{(pag. 42)}$$

Tali spazi non sono separabili, e non sono localmente compatti.

Se $(x^{(n)})_{n\in N}$ è una successione di punti di ℓ_∞ e $x^{(0)}$ è un punto di ℓ_∞ $[x^{(n)}=(x_k^{(n)})_{k\in N}\ \forall n\in N_0]$, la relazione:

$$\lim_{n\to\infty} x^{(n)} = x^{(0)} \quad \text{in } \ell_\infty$$

equivale al fatto che:

$$\lim_{n\to\infty} x_k^{(n)} = x_k^{(0)} \quad \text{uniformemente rispetto a } k\in N.$$

Se $(x_n)_{n\in N}$ è una successione di punti di $L^\infty(\Omega)$ e se x_0 è un punto di $L^\infty(\Omega)$, la relazione:

$$\lim x_n = x_0 \quad \text{in } L^\infty(\Omega)$$

equivale al fatto che:

> comunque si scelga per ogni $n\in N_0$ un rappresentante x_n^* della classe di funzioni x_n, esiste un sottoinsieme Ω_0 di Ω, di misura nulla, tale che si abbia:
> $$\lim x_n^*(t) = x_0^*(t) \quad \text{uniformemente in } \Omega-\Omega_0.$$

2) ℓ_p ed $L^p(\Omega)$ (p reale ≥ 1) sono spazi vettoriali normati, con le norme definite da:

$$\|x\|_{\ell_p} = \left(\sum_{k\in N} |x_k|^p\right)^{1/p} \quad [x=(x_k)_{k\in N}], \qquad \|x\|_{L^p(\Omega)} = \left(\int_\Omega |x(t)|^p dt\right)^{1/p} \quad (1)$$

Tali spazi sono separabili, e non sono localmente compatti.

Se $(x^{(n)})_{n\in N}$ è una successione di punti di ℓ_p e $x^{(0)}$ è un punto di ℓ_p $[x^{(n)}=(x_k^{(n)})_{k\in N}\ \forall n\in N_0]$, la relazione:

$$\lim x^{(n)} = x^{(0)} \quad \text{in } \ell_p$$

significa che:

$$\lim_{n\to\infty} \sum_{k\in N} |x_k^{(n)} - x_k^{(0)}|^p = 0.$$

Se $(x_n)_{n\in N}$ è una successione di punti di $L^p(\Omega)$ e se x_0 è un punto di $L^p(\Omega)$, la relazione:

$$\lim x_n = x_0 \quad \text{in } L^p(\Omega)$$

significa che:

$$\lim_{\Omega} \int |x_n(t) - x_0(t)|^p dt = 0.$$

Tale relazione si esprime dicendo che la successione di funzioni $(x_n)_{n\in N}$ converge in media di ordine p alla funzione x_0 [in media quadratica per $p=2$, semplicemente in media per $p=1$].

---

(1) Se non risulta $p\geq 1$ i secondi membri non verificano la proprietà triangolare della norma: è questo il motivo per cui gli spazi ℓ_p, $L^p(\Omega)$ si considerano per $p\geq 1$.

Se e solo se $p=2$ le norme sono generate da un prodotto scalare. Il prodotto scalare in ℓ_2 è definito da:

$$(x,y)_{\ell_2} = \sum_{k=1}^{\infty} x_k \overline{y}_k ,$$

e quello in $L^2(\Omega)$ è definito da:

$$(x,y)_{L^2(\Omega)} = \int_{\Omega} x(t)\overline{y(t)}\,dt .$$

In una Appendice al n.5, a pag.132.2, riportiamo la dimostrazione di alcuni risultati relativi allo spazio ℓ_p.

3) $\underline{L^p(\Omega,\mu)\ \ (1\le p\le+\infty)}$ è uno spazio vettoriale normato, con la norma definita da:

$$\|x\|_{L^\infty(\Omega,\mu)} = \sup_{t\in\Omega} |x(t)| \quad [1], \qquad \|x\|_{L^p(\Omega,\mu)} = \left(\int_{\Omega} |x(t)|^p d\mu \right)^{1/p} \quad (p\neq+\infty).$$

Il fatto che lo spazio, come $L^p(\Omega)$, sia separabile per $p\neq\infty$ e non sia localmente compatto, dipende dalla misura μ.

Se e solo se $p=2$ la norma in $L^p(\Omega,\mu)$ è generata dal prodotto scalare:

$$(x,y)_{L^2(\Omega,\mu)} = \int_{\Omega} x(t)\overline{y(t)}\,d\mu .$$

La convergenza in $L^\infty(\Omega,\mu)$ si caratterizza come nel caso che μ sia la misura di Lebesgue: naturalmente il sottoinsieme Ω_0 di Ω menzionato a pag.99 ha misura μ nulla, ossia $\mu(\Omega_0)=0$.

Il significato della convergenza in $L^p(\Omega,\mu)$ è analogo a quello di $L^p(\Omega)$.

4) $\underline{C^0(T) \text{ è uno spazio vettoriale normato}}$, con la norma definita da:

$$\|x\|_{C^0(T)} = \sup_{t\in T} |x(t)| .$$

Supposto che T sia uno pseudodominio, lo spazio è separabile se e solo se T è limitato, e non è localmente compatto.

Se $(x_m)_{m\in\mathbb{N}}$ è una successione di punti di $C^0(T)$ e x_0 è un punto di $C^0(T)$, la relazione:

$$\lim x_m = x_0 \quad \text{in } C^0(T)$$

equivale al fatto che:

$$\lim x_m(t) = x_0(t) \quad \text{uniformemente in } T.$$

Se T è compatto, lo spazio vettoriale costituito dalle restrizioni a T dei polinomi nel campo reale è denso in $C^0(T)$.

5) $\underline{C^m(T) \text{ è uno spazio vettoriale normato}}$, con la norma definita da:

$$\|x\|_{C^m(T)} = \sum_{|\alpha|=0}^{m} \sup_{t\in T} |D^\alpha x(t)| = \sum_{|\alpha|=0}^{m} \|D^\alpha x\|_{C^0(T)}$$

---

[1] Naturalmente nell'estremo superiore essenziale interviene la misura μ.

Si dimostra che se l'interno di T è dotato della proprietà di cono [1], sopprimendo gli addendi con $|\alpha|=1,\dots,m-1$ si ottiene una norma equivalente [naturalmente se $m>1$].

Lo spazio è separabile se e solo se T è limitato; non è localmente compatto.

Se $(x_n)_{n\in\mathbb{N}}$ è una successione di punti di $C^m(T)$ e se x_0 è un punto di $C^m(T)$, la relazione:

$$\lim x_n = x_0 \quad \text{in } C^m(T)$$

equivale al fatto che:

$$\lim D^\alpha x_n(t) = D^\alpha x_0(t) \text{ uniformemente in } T, \quad \forall \alpha : 0\le|\alpha|\le m.$$

6) $C^{0,\lambda}(T)$ è uno spazio vettoriale normato, con la norma definita da:

$$\|x\|_{C^{0,\lambda}(T)} = \|x\|_{C^0(T)} + [x]_\lambda .$$

Lo spazio non è separabile, nè localmente compatto.

Se $(x_n)_{n\in\mathbb{N}}$ è una successione di punti di $C^{0,\lambda}(T)$ e se x_0 è un punto di $C^{0,\lambda}(T)$, la relazione:

$$\lim x_n = x_0 \quad \text{in } C^{0,\lambda}(T)$$

equivale al fatto che:

$$\lim x_n(t) = x_0(t) \text{ uniformemente in } T, \quad \lim [x_n - x_0]_\lambda = 0 .$$

Detto X un sottoinsieme di $C^{0,\lambda}(T)$, dire che X è limitato in $C^{0,\lambda}(T)$ equivale a dire che le funzioni appartenenti ad X sono equilimitate (ossia, i loro moduli hanno un maggiorante comune), ed inoltre sono equi-höldeziane, ossia esiste $\kappa>0$ tale che:

$$|x(t')-x(t'')| \le \kappa |t'-t''|^\lambda \quad \forall t',t''\in T \text{ e } \forall x\in X. \tag{1}$$

La (1) implica che le funzioni appartenenti ad X sono equicontinue, ossia godono della proprietà:

$$\forall \varepsilon>0 \ \exists \delta>0 : \text{per } t',t''\in T, \ \left(|t'-t''|<\delta\right) \Longrightarrow \left(|x(t')-x(t'')|<\varepsilon \ \forall x\in X\right).$$

Un noto teorema di Ascoli-Arzelà assicura che un insieme di funzioni equicontinue ed equilimitate è precompatto in C^0. Pertanto un sottoinsieme X di $C^{0,\lambda}(T)$ limitato in $C^{0,\lambda}(T)$ è precompatto in $C^0(T)$.

7) $C^{m,\lambda}(T)$ è uno spazio vettoriale normato, con la norma definita da:

$$\|x\|_{C^{m,\lambda}(T)} = \|x\|_{C^m(T)} + \sum_{|\alpha|=m} [D^\alpha x]_\lambda$$

Si dimostra che se l'interno di T è dotato della proprietà di cono, tale norma è equivalente a quella definita da:

$$\|x\|_{C^{m,\lambda}(T)} = \|x\|_{C^0(T)} + \sum_{|\alpha|=m} [D^\alpha x]_\lambda$$

Lo spazio non è separabile, nè localmente compatto.

Se $(x_n)_{n\in\mathbb{N}}$ è una successione di punti di $C^{m,\lambda}(T)$ e se x_0 è un punto di $C^{m,\lambda}(T)$, la relazione:

$$\lim x_n = x_0 \quad \text{in } C^{m,\lambda}(T)$$

equivale al fatto che:

---

[1] Un aperto Ω di $\mathbb{R}^n$ si dice dotato della proprietà di cono quando esiste un cono C tale che sia verificata la condizione: per ogni $x\in\Omega$ esiste un cono di vertice x contenuto in Ω e congruente a C.

Per $n>1$, detti a e b due numeri positivi, si chiama cono di vertice $o=(0,\dots,0)$ l'insieme C_o dei punti $(x_1,\dots,x_n)\in\mathbb{R}^n$ tali che $x_1^2+\dots+x_{n-1}^2 \le a^2 x_n^2$, $0\le x_n\le b$; per $n=1$, l'intervallo $[0,a]$ con $a>0$ si chiama anche, per uniformità di linguaggio, un cono di vertice 0. Qualunque sia l'operatore lineare $L:\mathbb{R}^n\to\mathbb{R}^n$ che si identifichi con una matrice ortogonale di determinante 1, e qualunque sia $v\in\mathbb{R}^n$, l'insieme $C=L(C_o)+v$ si chiama un cono di vertice v.

Due coni C_1 e C_2 si dicono congruenti quando esistono un operatore lineare L del tipo predetto e un punto x_0 di $\mathbb{R}^n$ tali che $C_2 = L(C_1)+x_0$.

$$\lim x_n = x_0 \text{ in } C^m(T), \quad \lim \left[D^\alpha x_n - D^\alpha x_0\right]_\lambda = 0 \quad \forall \alpha : |\alpha| = m.$$

Se X è un sottoinsieme di $C^{m,\lambda}(T)$ limitato in $C^{m,\lambda'}(T)$, X è precompatto in $C^m(T)$.

8) $C^\infty(T)$ è uno spazio vettoriale topologico localmente convesso [1], munito di una topologia (che non riportiamo) per la quale una successione $(x_n)_{n\in N}$ converge ad x_0 se e solo se:

$$\forall \text{ compatto } K \subset T, \forall m \in N_0 \text{ si ha } \lim x_n = x_0 \text{ in } C^m(K)$$

Lo spazio non è separabile; è localmente compatto. E' uno spazio di Hausdorff, ed è metrizzabile.

Avvertiamo che, se Ω è un aperto di R^k, lo spazio $C^\infty(\Omega)$ viene spesso denotato con $\mathcal{E}(\Omega)$.

9) $C_c^0(\Omega)$ è uno spazio vettoriale topologico localmente convesso, munito di una topologia (che non riportiamo) per la quale una successione $(x_n)_{n\in N}$ converge ad x_0 se e solo se:

$$\exists \text{ un compatto } T \subset \Omega : \operatorname{supp} x_n \subseteq T \ \forall n \in N_0, \quad \lim x_n = x_0 \text{ in } C^0(T).$$

Lo spazio non è separabile e non è localmente compatto. E' uno spazio di Hausdorff, ma non è metrizzabile.

10) $C_c^m(\Omega)$ è uno spazio vettoriale topologico localmente convesso, munito di una topologia (che non riportiamo) per la quale una successione $(x_n)_{n\in N}$ converge ad x_0 se e solo se:

$$\exists \text{ un compatto } T \subset \Omega : \operatorname{supp} x_n \subseteq T \ \forall n \in N_0, \quad \lim x_n = x_0 \text{ in } C^m(T).$$

Lo spazio non è separabile e non è localmente compatto. E' uno spazio di Hausdorff, ma non è metrizzabile.

11) $C_c^\infty(\Omega)$ è uno spazio vettoriale topologico localmente convesso, munito di una topologia (che non riportiamo) per la quale una successione $(x_n)_{n\in N}$ converge ad x_0 se e solo se:

$$\exists \text{ un compatto } T \subset \Omega : \operatorname{supp} x_n \subseteq T \ \forall n \in N_0, \quad \lim x_n = x_0 \text{ in } C^\infty(T).$$

Come $C^\infty(\Omega)$ lo spazio non è separabile, ed è localmente compatto. E' uno spazio di Hausdorff, non è metrizzabile. $C_c^\infty(\Omega)$, sottospazio vettoriale di $C^\infty(\Omega)$, è denso in $C^\infty(\Omega)$ [Schwartz, pag. 89, 21, 88].

Qualunque sia il numero reale $p \geq 1$, $C_c^\infty(\Omega)$ è denso in $L^p(\Omega)$: naturalmente in tale affermazione si intende che ad ogni $x \in C_c^\infty(\Omega)$ sia stata aggiunta la classe delle funzioni q.o. uguali ad x.

Avvertiamo che lo spazio $C_c^\infty(\Omega)$ viene spesso denotato con $\mathcal{D}(\Omega)$.

12) $L^p_{loc}(\Omega)$ $(p \geq 1)$ è uno spazio vettoriale topologico localmente convesso, munito di una topologia (che non riportiamo) per la quale una successione $(x_n)_{n\in N}$ converge ad x_0 se e solo se:

$$\forall \text{ compatto } T \subset \Omega, \quad \lim x_n = x_0 \text{ in } L^p(T).$$

Lo spazio è separabile; non è localmente compatto.

---

[1] La proprietà di essere localmente convesso non è stata evidenziata per gli spazi precedenti, in quanto, ovviamente, ogni spazio vettoriale normato è localmente convesso.

13) $W^{m,p}(\Omega)$ $(p \geq 1)$ è uno spazio vettoriale normato, con la norma definita da:

$$\|x\|_{W^{m,p}(\Omega)} = \left(\sum_{0 \leq |\alpha| \leq m} \|D^\alpha x\|^p_{L^p(\Omega)} \right)^{1/p} \quad \left[\begin{array}{l} \text{per } p = +\infty \text{ eliminare} \\ \text{gli esponenti } p, \frac{1}{p} \end{array} \right]$$

Lo spazio è separabile per $p \neq +\infty$; non è localmente compatto.

Se $(x_n)_{n \in \mathbb{N}}$ è una successione di punti di $W^{m,p}(\Omega)$, e x_0 è un punto di $W^{m,p}(\Omega)$, la relazione $\lim x_n = x_0$ in $W^{m,p}(\Omega)$ equivale al fatto che:

$$\lim D^\alpha x_n = D^\alpha x_0 \text{ in } L^p(\Omega), \quad \forall \alpha : 0 \leq |\alpha| \leq m.$$

Se e solo se $p = 2$ la norma è generata dal prodotto scalare definito da:

$$(x,y)_{W^m(\Omega)} = \sum_{0 \leq |\alpha| \leq m} \int_\Omega D^\alpha x \, \overline{D^\alpha y} \, dt.$$

Lo spazio $W^{m,p}(\Omega)$ si chiama uno <u>spazio di Sobolev</u>.

14) <u>Lo spazio di Sobolev $H^{m,p}(\Omega)$ $(p \geq 1)$.</u>

L'insieme degli elementi di $W^{m,p}(\Omega)$ che hanno un rappresentante di classe $C^m(\bar{\Omega})$ costituisce un sottospazio vettoriale di $W^{m,p}(\Omega)$; la chiusura di tale sottospazio [in $W^{m,p}(\Omega)$] si denota con $H^{m,p}(\Omega)$, e per $p = 2$ con $H^m(\Omega)$.

Un elemento di $W^{m,p}(\Omega)$ appartiene ad $H^{m,p}(\Omega)$ quando, pur non avendo necessariamente un rappresentante di classe $C^m(\bar{\Omega})$, è il limite [nello spazio $W^{m,p}(\Omega)$] di una successione di elementi di $W^{m,p}(\Omega)$ ciascuno dei quali ha un rappresentante di classe $C^m(\bar{\Omega})$: tale asserzione sarà ovvia dopo la (1.7) di pag. 108.

Lo spazio $H^{m,p}(\Omega)$ si considera munito della stessa norma di $W^{m,p}(\Omega)$; di conseguenza per $p = 2$ la norma è generata dal prodotto scalare in $W^m(\Omega)$.

Se Ω è un intervallo di $\mathbb{R}$, lo spazio $H^m(\Omega)$ coincide con $W^m(\Omega)$.

Avvertiamo che da qualche tempo lo spazio $H^{m,p}(\Omega)$ si trova sempre più di frequente introdotto come la chiusura in $W^{m,p}(\Omega)$ del sottospazio costituito dagli elementi di $W^{m,p}(\Omega)$ che hanno un rappresentante di classe $C^m(\Omega)$ [quindi non necessariamente di classe $C^m(\bar{\Omega})$ come richiesto dianzi]. Con tale definizione, peraltro da noi non adottata, è stato dimostrato da Meyers e Serrin [cfr. ad es. Adams, Sobolev spaces (1975), (3.16) pag. 52] che lo spazio $H^{m,p}(\Omega)$ coincide con $W^{m,p}(\Omega)$. Intendendo $H^{m,p}(\Omega)$ secondo la definizione da noi introdotta, l'uguaglianza $H^{m,p}(\Omega) = W^{m,p}(\Omega)$ è stata dimostrata con opportune ipotesi sull'aperto Ω (in particolare, come già detto, nell'ipotesi che Ω sia un intervallo di $\mathbb{R}$).

15) <u>Lo spazio di Sobolev $H_0^{m,p}(\Omega)$ $(p \geq 1)$.</u>

L'insieme degli elementi di $W^{m,p}(\Omega)$ che hanno un rappresentante di classe $C_0^m(\Omega)$ è un sottospazio vettoriale di $H^{m,p}(\Omega)$; la chiusura di tale sottospazio, anch'essa sottospazio di $H^{m,p}(\Omega)$, si denota con $H_0^{m,p}(\Omega)$ o con $W_0^{m,p}(\Omega)$, e per $p = 2$ con $H_0^m(\Omega)$ o con $W_0^m(\Omega)$.

Un elemento di $H_0^{m,p}(\Omega)$ è dunque il limite [nello spazio $W^{m,p}(\Omega)$] di una successione di elementi di $W^{m,p}(\Omega)$ ciascuno dei quali ha un rappresentante di classe $C^m(\Omega)$ a supporto compatto in Ω; non è detto però che un elemento di $H_0^{m,p}(\Omega)$ sia a supporto compatto in Ω (*). In opportune ipotesi su Ω, si può dire che un elemento di $H_0^{m,p}(\Omega)$,

(*) La nozione di supporto [cfr. es. 14 pag. 57] si estende alle funzioni definite quasi ovunque, nel modo seguente. Detta x una funzione definita q.o. in Ω, consideriamo la famiglia $(\Omega_i)_{i \in I}$ degli aperti contenuti in Ω tali che $x = 0$ q.o. in Ω_i (supposto che ne esistano), e denotiamo con Ω_0 l'unione di tutti gli Ω_i. L'insieme $\overline{\Omega - \Omega_0}$ si chiama il <u>supporto di x</u>. Si dimostra che $x = 0$ q.o. in Ω_0 (non è banale, perché l'insieme I non è numerabile).

Se non esiste alcun aperto nel quale $x = 0$ q.o., il <u>supporto di x</u> è per definizione la chiusura di Ω.

quando sussiste un certo legame tra m, p e il numero k delle variabili, hanno "traccia nulla" sulla frontiera di Ω: la nozione di "traccia", di cui non possiamo occuparci, consente di esprimere, per così dire, il comportamento su $\partial\Omega$ degli elementi di $H_0^{m,p}(\Omega)$. Il significato di "traccia nulla su $\partial\Omega$" sarà chiarito tra breve.

Anche $H_0^{m,p}(\Omega)$ si considera munito della stessa norma di $W^{m,p}(\Omega)$ [es.13) pag.103], e quindi per $p=2$ la norma è generata dal prodotto scalare in $W^{m,p}(\Omega)$.

Si dimostra che il sottospazio $C_0^\infty(\Omega)$ di $H_0^{m,p}(\Omega)$ è denso in $H_0^{m,p}(\Omega)$ [naturalmente si intende che ad ogni $x \in C_0^\infty(\Omega)$ sia stata aggregata la classe delle funzioni uguali ad x q.o.].

Se Ω ha misura finita, la norma in $H_0^m(\Omega)$ è equivalente alla norma:

$$\|x\|_{H_0^m(\Omega)} = \left(\sum_{|\tau|=m} \int_\Omega |D^\tau x|^2 dt \right)^{\frac{1}{2}}$$

che è generata dal prodotto scalare:

$$(x,y)_{H_0^m(\Omega)} = \sum_{|\tau|=m} \int_\Omega D^\tau x \, D^\tau y \, dt \ .$$

In particolare, per $m=1$:

$$(2) \qquad \|x\|_{H_0^1(\Omega)} = \left(\int_\Omega |\nabla x|^2 dt \right)^{\frac{1}{2}} , \qquad (x,y)_{H_0^1(\Omega)} = \int_\Omega \nabla x \cdot \nabla y \, dt \ .$$

Per $m=1$ l'equivalenza si deduce dalla <u>disuguaglianza di Friedrics</u> (o <u>di Poincaré</u>): <u>Se l'aperto Ω ha misura finita, esiste una costante c tale che</u>:

$$(3) \qquad \int_\Omega |x(t)|^2 dt \le c \int_\Omega |\nabla x|^2 dt \quad \forall x \in H_0^1(\Omega) \ .$$

Nel caso $k=1$ e $\Omega =]a,b[$ la (3) sussiste con $c=(b-a)^2$ e quindi diviene:

$$(4) \qquad \int_a^b |x(t)|^2 dt \le (b-a)^2 \int_a^b |x'(t)|^2 dt \quad \forall x \in H_0^1(]a,b[) \ .$$

— — — — — — — — — —

Per dimostrare l'equivalenza tra la norma in $H_0^1(\Omega)$ e la norma definita dalla prima delle (2), diciamola $\|x\|_1$, essendo banale che $\|x\|_1 \le \|x\|_{H_0^1(\Omega)}$ basta osservare che per la (3) si ha:

$$\|x\|^2_{H_0^1(\Omega)} = \int_\Omega |x(t)|^2 dt + \int_\Omega |\nabla x|^2 dt \le c \|x\|_1^2 + \|x\|_1^2 = (c+1)\|x\|_1^2 \ .$$

Per $m>1$ l'equivalenza si stabilisce analogamente dopo aver osservato che, detto x un elemento di $H_0^m(\Omega)$, per ciascun multiindice α tale che $|\alpha| \in \{1, \dots, m-1\}$ risulta $D^\alpha x \in H_0^1(\Omega)$ e quindi è applicabile la (3) con $D^\alpha x$ in luogo di x. Occorre poi tener presente che per ogni $h \in \{1, \dots, m-1\}$ le derivate deboli di una derivata debole di ordine h, sono derivate deboli di ordine $h+1$.

— — — — — — — — — — —

Con riferimento al caso particolare $k=1$, $p=2$, $\Omega =]a,b[$ vogliamo ora enunciare uno dei <u>teoremi di Sobolev</u>, concernenti gli spazi $H^{m,p}(\Omega)$, $H_0^{m,p}(\Omega)$ con k, m, p qualunque ed Ω aperto di R^k: in certe ipotesi su Ω, e sussistendo opportuni legami tra i numeri k, m, p, tali teoremi assicurano che gli elementi dei predetti spazi appartengono anche a determinati spazi "di migliore qualità"; in alcuni casi hanno un rappresentante di classe $C^r(\bar\Omega)$ o di classe $C^{r,\lambda}(\bar\Omega)$, con precisati valori di r e λ: ad esempio, in certe ipotesi su Ω si può dire, nel caso di due variabili, che gli elementi di $H^2(\Omega)$ hanno un rappresentante continuo in $\bar\Omega$.

Nei casi in cui ogni $x \in H^{m,p}(\Omega)$ ha un rappresentante $\tilde x$ di classe $C^r(\bar\Omega)$, viene poi dimostrato che x appartiene ad $H_0^{m,p}(\Omega)$ se e solo se <u>il rappresentante $\tilde x$ si annulla su $\partial\Omega$ insieme alle sue derivate (quando $r>1$) fino a quelle di ordine $r-1$</u>: nei casi anzidetti è questa la proprietà degli elementi di $H_0^{m,p}(\Omega)$, prima accennata, di avere <u>traccia nulla su $\partial\Omega$</u>.

Con le considerazioni svolte abbiamo voluto dare un'idea dei teoremi di Sobolev. Quello che vogliamo qui enunciare, che sarà utilizzato nel seguito, è il seguente:

TEOREMA DI SOBOLEV. Sia $\Omega =]a,b[$ un intervallo aperto e limitato di R, e sia m un intero positivo.

Ogni elemento u di $H^m(\Omega)$ ha un rappresentante $u^* \in C^{m-1,\frac{1}{2}}(\bar{\Omega})$, ed esiste una costante c tale che:

$$\|u^*\|_{C^{m-1,\frac{1}{2}}(\bar{\Omega})} \leq c \|u\|_{H^m(\Omega)} \qquad \forall u \in H^m(\Omega) . \tag{5}$$

L'elemento u di $H^m(\Omega)$ appartiene al sottospazio $H_0^m(\Omega)$ se e solo se il rappresentante u^* si annulla negli estremi dell'intervallo Ω insieme alle sue derivate fino a quella di ordine $m-1$.

Per evidenziare l'utilità della (5), dimostriamo che se $u \in H_0^m(\Omega)$ il rappresentante u^* si annulla nei punti a e b insieme alle sue derivate di ordine $\leq m-1$.

Infatti se $u \in H_0^m(\Omega)$, u è il limite [in $W^m(\Omega)$, o equivalentemente in $H^m(\Omega)$] di una successione di elementi di $C_0^m(\Omega)$. Esiste cioè almeno una successione $(u_n)_{n\in N}$ di elementi di $C_0^m(\Omega)$ tale che:

$$u = \lim u_m \quad \text{in } H^m(\Omega),$$

ossia tale che:

$$\lim \|u - u_m\|_{H^m(\Omega)} = 0 . \tag{6}$$

Naturalmente nella (6) u_m è inteso come l'elemento di $H^m(\Omega)$ che ha u_m come rappresentante, e tale rappresentante è ovviamente anche elemento di $C^{m-1,\frac{1}{2}}(\bar{\Omega})$ (1). Detto u^* il rappresentante di u di classe $C^{m-1,\frac{1}{2}}(\bar{\Omega})$, dalla (5) con $u-u_m$ in luogo di u:

$$\|u^* - u_m\|_{C^{m-1,\frac{1}{2}}(\bar{\Omega})} \leq c \|u - u_m\|_{H^m(\Omega)}$$

consegue che $u^* = \lim u_m$ nello spazio $C^{m-1,\frac{1}{2}}(\bar{\Omega})$, e quindi che risulta:

$$\lim u_m(t) = u^*(t) , \quad \lim u_m'(t) = u^{*\prime}(t), \ \ldots, \ \lim u_m^{(m-1)}(t) = u^{*(m-1)}(t) \tag{7}$$

uniformemente in $[a,b]$. Essendo le funzioni u_m a supporto compatto in $]a,b[$, ciascuna di esse è identicamente nulla fuori di un compatto contenuto in $]a,b[$, sicché nei punti a e b ciascuna delle u_m si annulla insieme alle derivate di qualunque ordine. In virtù della (7), la funzione u^* si annulla dunque nei punti a e b insieme alle sue derivate fino a quella di ordine $m-1$.

L'enunciato del teorema di Sobolev in generale è riportato nell'Appendice [cfr. pag. 304].

(1) La derivata di ordine $m-1$ della funzione u_m è lipschitziana (in quanto ha derivata limitata), e quindi è hölderiana con qualunque esponente.

Capitolo V - Successioni e serie . Spazi metrici completi . Spazi di Banach. Spazi di Hilbert.

1. Successioni convergenti in uno spazio topologico.

Siano : S uno spazio topologico , $(x_n)_{n\in N}$ una successione di punti di S , x_0 un punto di S .

Come abbiamo già visto nel cap. I, n°7, si dice che la successione $(x_n)_{n\in N}$ converge ad x_0 quando è vera la proposizione seguente:

(1) Per ciascun intorno I di x_0, esiste $\nu\in N$ tale che:

$$x_n \in I \quad \forall n>\nu.$$

Osserviamo che:

(1.1) Se S è uno spazio metrico , e d è la metrica in S , la (1) equivale alla proposizione [1]:

(2) Per ciascun $\varepsilon>0$ esiste $\nu\in N$ tale che:

$$d(x_n,x_0)<\varepsilon \quad \forall n>\nu .$$

Poiché uno spazio metrico è uno spazio di Hausdorff, ogni successione convergente di punti di S ha un unico limite.[2]

(1.2) Se S è uno spazio vettoriale topologico, la (1) equivale alla proposizione [3]:

(3) Per ciascun intorno I di 0, esiste $\nu\in N$ tale che :

$$x_n-x_0 \in I \quad \forall n>\nu.$$

(1.3) Se S è uno spazio vettoriale normato , la (1) equivale alla proposizione [4]:

(4) Per ciascun $\varepsilon>0$ esiste $\nu\in N$ tale che:

$$\|x_n-x_0\|<\varepsilon \quad \forall n>\nu.$$

In virtù della (1.1), ogni successione convergente di punti di S ha un unico limite.

---

(1) Cfr. (1.2) pag.78.
(2) Cfr. (7.1) pag.19.
(3) Cfr. (2.1) pag.82 , equivalenza(1).
(4) Cfr. n.3 pag.88, eguaglianza (1), tenendo presente la (1.1) sopra enunciata.

(1.4) Se S é uno spazio pre-hilbertiano, la (1) equivale alla (4), dove naturalmente si intende che $\|x\|=\sqrt{(x,x)}$ $\forall x \in S$. Inoltre sussiste l'implicazione:

$$(5) \qquad \lim_{n\to\infty} x_n = x_0 \implies \left[\lim_{n\to\infty} (x_n,x)=(x_0,x) \quad \forall x \in S \right] .$$

L'implicazione opposta sussiste se e solo se S ha dimensione finita.

Quando si ha $\lim_{n\to\infty}(x_n,x)=(x_0,x)$ $\forall x \in S$, si dice che la successione $(x_n)_{n\in N}$ converge ad x_0 debolmente in S.

Dim. (5). Basta osservare che, in virtù della disuguaglianza di Schwarz [cfr.(5) del n.4, pag. 97], per ogni $n \in N$ e per ogni $x \in S$ si ha:

$$|(x_n,x)-(x_0,x)| = |(x_n-x_0,x)| \leq \|x_n-x_0\| \cdot \|x\| .$$

Della seconda parte dell'enunciato, ci limitiamo a dimostrare che se S ha dimensione finita k, nella (5) vale l'implicazione opposta. Denotiamo allora con $\{e_1, \dots, e_k\}$ una base di S ortonormale, ossia tale che:

$$(e_i,e_j)=\begin{cases} 0 & \text{se } i \neq j \\ 1 & \text{se } i=j \end{cases} \quad (1)$$

Posto:

$$x_n=\sum_{r=1}^{k} \lambda_{nr} e_r \quad \forall n \in N \qquad x_0=\sum_{r=1}^{k} \lambda_{0r} e_r$$

per ogni $n \in N$ e $\forall i \in \{1,\dots,k\}$ risulta:

$$(x_n,e_i)=\sum_{r=1}^{k} \lambda_{nr}(e_r,e_i)=\lambda_{ni}$$

$$(x_0,e_i)=\sum_{r=1}^{k} \lambda_{0r}(e_r,e_i)=\lambda_{0i} .$$

Pertanto, se vale quanto scritto a secondo membro di (5), risulta:

$$(6) \qquad \lim_{n\to\infty} \lambda_{ni}=\lambda_{0i} \qquad \forall i \in \{1,\dots,k\}$$

Poichè:

$$\|x_n-x_0\|=\Big\| \sum_{i=1}^{k} (\lambda_{ni}-\lambda_{0i})e_i \Big\| \leq \sum_{i=1}^{k} |\lambda_{ni}-\lambda_{0i}| \cdot \|e_i\| = \sum_{i=1}^{k} |\lambda_{ni}-\lambda_{0i}|$$

dalla (6) consegue:

$$\lim_{n\to\infty} \|x_n-x_0\| = 0 .$$

---

(1) Non dimostriamo l'esistenza di una siffatta base di S, in quanto nel seguito proveremo una proposizione più generale.

Dalla (2.5) di pag. 86 consegue:

(1.5) Sia S uno spazio vettoriale topologico di Hausdorff, e siano: $(x_n)_{n\in N}$, $(y_n)_{n\in N}$ successioni di punti di S , x_0, y_0 punti di S , λ elemento di K.

Si ha:

$$\lim_{n\to\infty} x_n = x_0 \implies \lim_{n\to\infty} \lambda x_n = \lambda x_0$$

$$\lim_{n\to\infty} x_n = x_0 \ \text{e} \ \lim_{n\to\infty} y_n = y_0 \implies \lim_{n\to\infty} (x_n + y_n) = (x_0 + y_0).$$

Di immediata verifica sono le proposizioni:

(1.6) Se S è uno spazio metrico (1), ogni successione convergente di punti di S è limitata.

Dim. Se la successione $(x_n)_{n\in N}$ converge ad x_0, la successione $(d(x_n, x_0))_{n\in N}$ di numeri reali è infinitesima, e quindi è limitata. Pertanto esiste $r>0$ tale che $d(x_n, x_0) \leq r$ $\forall n \in N$. Ciò vuol dire che l'insieme degli elementi della successione $(x_n)_{n\in N}$ è contenuto nel cerchio $C[x_0, r]$.

(1.7). Sia X un sottoinsieme di uno spazio topologico S.
Se un punto x_0 è limite di una successione di punti di X, x_0 appartiene alla chiusura di X.
Viceversa, se S è uno spazio metrico, ogni punto della chiusura di X è il limite di una successione di punti di X.

Dim. La prima parte dell'enunciato è ovvia. Nell'ipotesi che S sia uno spazio metrico, denotiamo con x_0 un punto della chiusura di X. Per ogni $n\in N$ esiste allora almeno un punto x_n di X appartenente al cerchio aperto $C(x_0, \frac{1}{n})$; risultando:

$$d(x_n, x_0) < \frac{1}{n},$$

la successione $(x_n)_{n\in N}$ converge ad x_0.

Dalla prima parte della (1.7) consegue che se X è un sottoinsieme chiuso di uno spazio topologico, ogni successione convergente di punti di X ha per limite un punto di X; tale proprietà è caratteristica per gli insiemi chiusi di uno spazio metrico. Sussiste infatti il teorema:

---

(1) Quindi anche se S è uno spazio vettoriale normato, o uno spazio pre-hilbertiano.

(1.8) Sia X un sottoinsieme di uno spazio metrico S.

L'insieme X è chiuso se e solo se ogni successione convergente di punti di X ha per limite un punto di X.

Dim. Per quanto dianzi osservato, si tratta di far vedere che se ogni successione convergente di punti di X ha per limite un punto di X, l'insieme X è chiuso.

Sia $x_0 \in \overline{X}$: per la (1.7), x_0 è il limite di una successione di punti di X e quindi, per l'ipotesi posta, x_0 appartiene ad X. Ciò prova che $\overline{X} \subseteq X$, e pertanto $X = \overline{X}$, ossia X è chiuso.

Una conseguenza della (1.8) è l'osservazione seguente.

(1.9) Sia Y un sottoinsieme di uno spazio pre-hilbertiano S.

Il complemento ortogonale $Y^\perp$ di Y [1]) è un sottospazio vettoriale di S, chiuso in S.

Dim. Essendo ovvio che $Y^\perp$ è un sottospazio vettoriale di S, proviamo che $Y^\perp$ è chiuso.

Qualunque sia la successione convergente $(x_n)_{n\in\mathbb{N}}$ di punti di $Y^\perp$, detto x_0 il suo limite si ha, in virtù della (5) di pag. 107:

$$\lim_{n\to\infty} (x_n, x) = (x_0, x) \qquad \forall x \in S.$$

Di conseguenza per ogni $x \in Y$, essendo $(x_n, x) = 0$ $\forall n \in \mathbb{N}$ in quanto $x_n \in Y^\perp$, risulta $(x_0, x) = 0$.

Dunque x_0 è ortogonale ad ogni elemento di Y, ossia $x_0 \in Y^\perp$. Dopo ciò, in virtù della (1.8), resta provato che $Y^\perp$ è chiuso.

Consideriamo ora una funzione f definita in un sottoinsieme X di uno spazio topologico S, a valori in uno spazio topologico S'.

Si verifica subito che, se f è continua in un punto $x_0 \in X$, per ogni successione $(x_n)_{n\in\mathbb{N}}$ di punti di X convergente ad x_0 si ha:

$$\lim_{n\to\infty} f(x_n) = f(x_0)$$

Dim. Qualunque sia I' intorno di $f(x_0)$, esiste un intorno I di x_0 tale che:

$$f(x) \in I' \quad \forall x \in I;$$

d'altra parte, per la convergenza di $(x_n)_{n\in\mathbb{N}}$ ad x_0, esiste $\nu \in \mathbb{N}$ tale che:

$$x_n \in I \quad \forall n > \nu.$$

Pertanto risulta $f(x_n) \in I'$ $\forall n > \nu$, e l'asserto è provato.

La precedente osservazione può esprimersi dicendo che una funzione continua in un punto è anche "continua per successioni" nel punto. Se l'insieme di definizione è un sottoinsieme di uno spazio metrico, il risultato si inverte. Precisamente sussiste il teorema:

(1.10) Siano: X un sottoinsieme di uno spazio metrico S, S' uno spazio topologico, f una funzione definita in X ed a valori in S', x_0 un punto di X.

Le seguenti proposizioni sono equivalenti:

a) f è continua in x_0.

b) Per ogni successione $(x_n)_{n\in N}$ di punti di X convergente ad x_0 risulta:

$$\lim_{n\to\infty} f(x_n) = f(x_0).$$

Dim. Si tratta di far vedere che b) $\Rightarrow$ a), e a tale scopo, procedendo per assurdo, neghiamo la a). Esiste allora un intorno I' di $f(x_0)$ godente della proprietà: per ogni intorno I di x_0, esiste almeno un punto $x \in I \cap X$ tale che $f(x) \notin I'$. Ciò implica che per ogni $n\in N$ esiste almeno un punto $x_n \in X \cap C(x_0, \frac{1}{n})$ tale che $f(x_n) \notin I'$: pertanto la successione $(f(x_n))_{n\in N}$ non converge ad $f(x_0)$. Ma ciò contraddice la b) in quanto, detta d la metrica per S, risultando $d(x_n, x_0) < \frac{1}{n}$ $\forall n \in N$, la successione $(x_n)_{n\in N}$ converge ad x_0.

Dimostriamo infine il seguente teorema:

(1.11) Siano: S uno spazio vettoriale normato, H un sottospazio vettoriale chiuso di S, $(X_n)_{n\in N}$ una successione di punti di S/H, X_0 un punto di S/H.

Se risulta:

$$\lim X_n = X_0 , \tag{7}$$

allora comunque si fissi $x_0 \in X_0$ esiste, per ogni $n\in N$, un punto $x_n \in X_n$ tale da aversi:

$$\lim x_n = x_0 . \tag{8}$$

Viceversa se esistono $x_0 \in X_0$ e, per ogni $n\in N$, un punto $x_n \in X_n$ tale da verificare la (8), allora vale la (7).

Dim. Ricordando la norma nello spazio quoziente S/H [cfr. pag. 93], per ogni $n\in N$ si ha:

$$\| X_n - X_0 \|_{S/H} = \inf_{y \in X_n - X_0} \|y\|_S$$

e quindi esiste $y_n \in X_n - X_0$ tale che:

$$\|y_n\|_S < \| X_n - X_0 \|_{S/H} + \frac{1}{n} .$$

Pertanto se vale la (7), la successione $(y_n)_{n\in N}$ è tale da aversi:

$$\lim \|y_n\|_S = 0 . \tag{9}$$

Ora comunque si fissi $x_0 \in X_0$, per ogni $n\in N$ esiste $x_n \in X$ tale che:

$$y_n = x_n - x_0 \qquad (1),$$

e dalla (9) consegue la (8).

Viceversa, se esistono $x_0 \in X_0$ e $x_n \in X_n$ $\forall n \in N$ tali da verificare la (8), dalla relazione:

$$\| X_n - X_0 \|_{S/H} \leq \| x_n - x_0 \|_S \quad \forall n \in N$$

consegue la (7).

---

2. Sottoinsiemi sequenzialmente compatti di uno spazio topologico.

Siano: S uno spazio topologico, X un sottoinsieme di S.
Si dice che X è sequenzialmente precompatto, quando da ogni successione di punti di X si può estrarre una successione convergente.

Si dice che X è sequenzialmente compatto, quando da ogni successione di punti di X si può estrarre una successione convergente ad un punto di X.

Sussiste il seguente teorema, utilissimo per le applicazioni:

(2.1) Se S è uno spazio metrico per ogni $X \subseteq S$ si ha che (2):

X sequenzialmente precompatto $\iff$ X precompatto

X sequenzialmente compatto $\iff$ X compatto .

Dim. Si vedano le successive pagine 111 bis e 111 ter.

---

(1) Il vettore $X_n - X_0$ dello spazio S/H è la differenza dei vettori X_n, X_0 e quindi è del tipo $z_n - x_0 + H$ con $z_n \in X_n$. Siccome $y_n \in X_n - X_0$, esiste $h_n \in H$ tale che:

$$(*) \qquad y_n = z_n - x_0 + h_n = (z_n + h_n) - x_0 ,$$

e $z_n + h_n$ appartiene ad X_n, in quanto $z_n \in X_n$. Posto $z_n + h_n = x_n$, dalla (*) segue allora $y_n = x_n - x_0$ con $x_n \in X_n$.

Cogliamo l'occasione per evidenziare che, se X ed Y sono due vettori dello spazio vettoriale S/H, il simbolo:

$$X - Y$$

da una parte denota la differenza U dei due vettori, ossia la classe di equivalenza $x_0 - y_0 + H$ con $x_0 \in X$ ed $y_0 \in Y$, dall'altra denota l'insieme costituito dai vettori di S ciascuno dei quali è la differenza di un vettore appartenente ad X e di un vettore appartenente ad Y. Si verifica subito che, detto S_0 quest'ultimo insieme, si ha $S_0 = U$ e quindi non possono sorgere equivoci. Infatti si ha:

$$z \in S_0 \Rightarrow \exists x \in X \text{ e } \exists y \in Y : z = x - y \Rightarrow z = (x_0 + h_0) - (y_0 + h_0') \text{ con } h_0, h_0' \in H \Rightarrow z \in (x_0 - y_0) + H \Rightarrow z \in U ;$$

$$z \in U \Rightarrow z \in (x_0 - y_0) + H \Rightarrow z = x_0 - y_0 + h \text{ con } h \in H \Rightarrow z = (x_0 + h) - y_0 = x - y_0 \text{ con } x \in X \Rightarrow z \in S_0 .$$

(2) Si ricordino le definizioni di pag. 14-15.

La dimostrazione del teor. (2.1) è alquanto complessa: accenneremo a quella della seconda equivalenza, supponendo X infinito essendo banale l'asserto per X finito.

Nell'ipotesi che X sia sequenzialmente compatto, si dimostrano successivamente le affermazioni che seguono:

a) Per ogni $\varepsilon > 0$ esiste un ricoprimento finito di X costituito da cerchi aperti di raggio ε con i centri in punti di X.

b) Esiste un sottoinsieme numerabile di X denso in X.

c) Sia $(y_n)_{n\in N}$ una successione di punti di X avente codominio denso in X [esistente per b)]. Detto $\mathcal{R}$ un arbitrario ricoprimento aperto di X, sia $\mathcal{C} = \mathcal{C}(\mathcal{R})$ l'insieme costituito da quelli, fra i cerchi aperti con i centri nei punti y_n, il cui raggio sia del tipo $1/m$ con $m \in N$ e che siano contenuti in almeno un elemento di $\mathcal{R}$.
L'insieme $\mathcal{C}$, ovviamente numerabile, è un ricoprimento di X.

d) Sia $(A_n)_{n\in N}$ una successione di cerchi aperti il cui codominio è il ricoprimento $\mathcal{C}$ costruito in c).
Esiste un numero finito di termini della successione $(A_n)_{n\in N}$ che costituiscono un ricoprimento $\mathcal{C}'$ di X.

e) Per ogni elemento Γ del ricoprimento $\mathcal{C}'$ menzionato in d), che è finito, si scelga un elemento di $\mathcal{R}$ contenente Γ, esistente perché $\mathcal{C}' \subset \mathcal{C}$ [si ricordi la definizione di $\mathcal{C}$ effettuata in c)], e si denoti con $\mathcal{R}'$ l'insieme dei predetti elementi di $\mathcal{R}$.
$\mathcal{R}'$ è finito, è contenuto in $\mathcal{R}$, ed è un ricoprimento di X.

Dim. a). Scelto $x_1 \in X$, sia Γ_1 il cerchio aperto di centro x_1 e raggio ε. Se $\Gamma_1 \supseteq X$ l'asserto è provato, altrimenti esiste $x_2 \in X - \Gamma_1$: sia Γ_2 il cerchio aperto di centro x_2 e raggio ε. Se $\{\Gamma_1, \Gamma_2\}$ è un ricoprimento di X, l'asserto è provato, altrimenti esiste $x_3 \in X - (\Gamma_1 \cup \Gamma_2)$. Così continuando il procedimento si arresta (e ciò prova l'asserto). Infatti in caso contrario si viene a costruire una successione $(x_n)_{n\in N}$ di punti di X, e da questa per l'ipotesi si può estrarre una successione convergente $(x_{k_n})_{n\in N}$. Essendo quest'ultima una successione di Cauchy, devono esistere due suoi termini consecutivi $x_{k_\nu}, x_{k_{\nu+1}}$ aventi distanza minore di ε: ciò è assurdo, perché $x_{k_{\nu+1}}$ per costruzione non appartiene all'unione dei cerchi $\Gamma_1, \dots, \Gamma_{k_{\nu+1}-1}$ (fra i quali c'è Γ_{k_ν}) e perciò $d(x_{k_\nu}, x_{k_{\nu+1}}) \geq \varepsilon$.

Dim b). Comunque si fissi $p \in N$, per la a) esiste un ricoprimento finito di X, diciamolo $\mathcal{R}_p$, costituito da cerchi aperti di raggio $1/p$ con i centri in punti di X. Detto Y_p l'insieme dei centri dei cerchi che costituiscono $\mathcal{R}_p$, l'insieme $Y = \bigcup_{p\in N} Y_p$ è ovviamente numerabile, e si ha $\overline{Y} = X$. Infatti, se x_0 è un punto di X e Γ è un cerchio aperto di centro x_0 e raggio $\varepsilon > 0$, per ogni $p \in N$ il punto x_0 deve appartenere ad almeno uno dei cerchi che costituiscono $\mathcal{R}_p$, sia esso Γ_p di centro c_p; scelto p tale che $\frac{1}{p} < \varepsilon$, si ha $d(x_0, c_p) < \varepsilon$ e quindi $c_p \in \Gamma$. Siccome $c_p \in Y$, ciò prova che x_0 è aderente ad Y.

Dim c). Sia x_0 un punto di X, e sia B_0 un elemento di $\mathcal{R}$ cui appartenga x_0. Detto Γ_0 un cerchio aperto di centro x_0 contenuto in B_0 e detto ρ il suo raggio, sia $m \in N$ tale che $\frac{1}{m} < \frac{\rho}{2}$, e sia $n \in N$ tale che y_n appartenga al cerchio di centro x_0 e raggio $1/m$. Allora x_0 appartiene al cerchio di centro y_n e raggio $1/m$, che è un elemento di $\mathcal{C}$ in quanto è contenuto in $\Gamma_0 \subseteq B_0 \in \mathcal{R}$.

Dim. d). Neghiamo la tesi. Allora per ogni $n \in N$ esiste almeno un punto x_n di X che non appartiene all'unione dei cerchi $A_1, \dots, A_n$. Per l'ipotesi, dalla successione $(x_n)_{n\in N}$ si può

estrarre una successione $(x_{k_n})_{n\in N}$ convergente ad un punto $\bar{x}\in X$. Esiste allora $h\in N$ tale che $\bar{x}\in A_h$, e per la definizione di limite esiste $\nu > h$ tale che $x_\nu \in A_h$: ciò è assurdo per la costruzione di x_ν.

Dim. e). $\mathcal{R}'$ è un ricoprimento di X perché l'unione degli elementi di $\mathcal{R}'$ contiene l'unione degli elementi di $\mathcal{C}'$, e $\mathcal{C}'$ è un ricoprimento di X.

Viceversa, supponiamo X compatto.

Negando la tesi, ammettiamo che X non sia sequenzialmente compatto: allora esiste una successione $(x_n)_{n\in N}$ di punti di X tale che ogni estratta non converge ad un punto di X.

Il codominio Y della successione $(x_n)_{n\in N}$ è costituito da punti isolati [1], e perciò per ogni $n\in N$ l'insieme $Y_n = \bar{Y} - \{x_n\}$ è chiuso [2].

Posto $A_n = S - Y_n$, la successione $(A_n)_{n\in N}$ è un ricoprimento aperto di X [3], e quindi per l'ipotesi contiene un ricoprimento finito di X, sia esso $\{A_{n_1}, \dots, A_{n_h}\}$.

Scelto $m\in N$ tale che x_m sia diverso da $x_{n_1}, x_{n_2}, \dots, x_{n_h}$ [4], x_m appartiene ad almeno uno degli insiemi $A_{n_1}, \dots, A_{n_h}$ (poiché questi costituiscono un ricoprimento di X), sia esso A_{n_p}: ma ciò è assurdo, in quanto $x_m \in Y_{n_p}$ e quindi $x_m \notin A_{n_p}$.

---

[1] In caso contrario, si potrebbe costruire una estratta dalla $(x_n)_{n\in N}$ convergente ad un punto di X.

[2] Detto I_n un intorno aperto di x_n cui non appartenga alcun punto di Y, l'insieme $\bar{Y} - I_n$ è chiuso, e coincide con Y_n.

[3] Sia $\bar{x}$ un punto di X.

Se $\bar{x}\in Y$, detto n un elemento di N tale che $\bar{x} = x_n$, il punto $\bar{x}$ non appartiene ad Y_n, ossia appartiene ad $A_n = S - Y_n$.

Se $\bar{x}\notin Y$, $\bar{x}$ non può appartenere ad $\bar{Y}$ (altrimenti sarebbe di accumulazione per Y) e quindi non appartiene ad alcuno degli insiemi Y_n, sicché $\bar{x}$ appartiene addirittura a tutti gli A_n.

[4] Esiste certamente un siffatto x_m, altrimenti ci sarebbe un elemento della successione $(x_n)_{n\in N}$ che si ripete infinite volte, quindi esisterebbe una estratta convergente ad un punto di X.

3 . Spazi metrici completi . Spazi di Banach. Spazi di Hilbert.

Sia (S,d) uno spazio metrico, e sia $(x_n)_{n\in N}$ una successione di punti di S .

Si dice che $(x_n)_{n\in N}$ é una successione di Cauchy di (S,d) , quando soddisfa alla seguente condizione:

Per ciascun $\varepsilon>0$ esiste $\nu\in N$ tale che:

$$d(x_p,x_q)<\varepsilon \quad \forall p\geq\nu \text{ e } \forall q\geq\nu .$$

Evidentemente :

(3.1) Ogni successione convergente é una successione di Cauchy .

Nello spazio K^n (munito della metrica usuale), si ha reciprocamente che ogni successione di Cauchy é convergente (criterio di convergenza di Cauchy). Però tale asserzione non é vera in uno spazio metrico qualunque : esistono cioé degli spazi metrici nei quali non tutte le successioni di Cauchy sono convergenti.

Quanto sopra giustifica la seguente definizione . Si dice che lo spazio metrico (S,d) é completo , quando ogni successione di Cauchy di (S,d) é convergente.

Con riferimento ad un qualunque spazio metrico (S,d) , sia X un sottoinsieme di S .La restrizione di d ad $X\times X$, diciamola d_X ,é ovviamente una metrica per X ,e la topologia per X da essa generata é la topologia indotta da τ_d su X [essendo τ_d la topologia per S generata da d]. Perciò lo spazio metrico (X,d_X) si chiama sottospazio (metrico) di (S,d), di sostegno X.

Ciò premesso sussiste la proposizione seguente:

(3.2) Sia X un sottoinsieme di uno spazio metrico S .

Se il sottospazio di sostegno X é completo, allora X é chiuso . Viceversa, nell'ipotesi che S sia completo, se X é chiuso il sottospazio di sostegno X é anch'esso completo .

Dim. Siano d la metrica di S e d_X la metrica del sottospazio di sostegno X.

Supposto (X,d_X) completo, sia $(x_n)_{n\in N}$ una successione di punti di X convergente in S. La $(x_n)_{n\in N}$ é una successione di Cauchy di S [per la (3.1)] , e quindi anche di (X,d_X): ne segue che tale successione é convergente anche in (X,d_X) , e perciò il suo limite appartiene ad X . In virtù della (1.8) [pag.109] ne segue che X é chiuso.

Nell'ipotesi S completo, supponiamo viceversa che X sia chiuso. Detta $(x_n)_{n\in N}$ una successione di Cauchy di (X,d_X), e quindi anche di S , tale successione é convergente in S . Essendo X chiuso, ancora per la (1.8) il limite appartiene ad X , e quindi la successione converge anche in (X,d_X). Pertanto (X,d_X) é completo.

La nozione di completezza si può riferire , naturalmente , anche agli spazi vettoriali normati, nonché agli spazi pre-hilbertiani .
Se uno spazio vettoriale normato é completo, esso si chiama uno spazio di Banach; se uno spazio pre-hilbertiano é completo, esso si chiama uno spazio di Hilbert, oppure uno spazio euclideo. Poiché gli spazi pre-hilbertiani si possono considerare come particolari spazi vettoriali normati, possiamo dire che gli spazi di Hilbert sono particolari spazi di Banach (precisamente quelli la cui norma é definita mediante un prodotto scalare).

Dimostriamo ora che:

(3.3) Lo spazio prodotto di una famiglia finita di spazi di Banach e' anch'esso uno spazio di Banach.

— — — — — — — — —

Dim. Sia S lo spazio prodotto della famiglia $(S_i)_{i\in\{1,\dots,K\}}$ di spazi di Banach.
Per ogni successione di Cauchy di punti di S, diciamola $(x_n)_{n\in N}$, posto:

$$x_n = (x_{n1}, x_{n2}, \dots, x_{nK}) \quad \forall n\in N,$$

la successione $(x_{ni})_{n\in N}$ e' una successione di Cauchy di S_i, qualunque sia $i\in\{1,\dots,K\}$. Dopo ciò basta osservare che la convergenza della successione $(x_n)_{n\in N}$ equivale alla convergenza in S_i, per ogni i, della successione $(x_{ni})_{n\in N}$.

— — — — — — — — —

Anche lo spazio quoziente di uno spazio di Banach, rispetto ad un sottospazio chiuso [cfr. pag. 93], e' a sua volta uno spazio di Banach. Sussiste infatti il teorema:

(3.4) Siano: S uno spazio vettoriale normato, H un sottospazio vettoriale di S, chiuso in S. Vale l'implicazione:

$$S \text{ spazio di Banach} \Longrightarrow S/H \text{ spazio di Banach}.$$

Dim. Sia $(X_n)_{n\in N}$ una successione di Cauchy di S/H. Esiste allora una successione estratta da $(X_n)_{n\in N}$, diciamola $(X_{K_n})_{n\in N}$, tale che:

$$\|X_{K_{n+1}} - X_{K_n}\| < 1/n^2 \quad \forall n\in N.$$

Conseguentemente, per ogni $n\in N$ esiste $y_n \in X_{K_{n+1}} - X_{K_n}$ tale che:

$$\|y_n\|_S < 1/n^2.$$

Posto $z_1 = y_1$, nonché:

(1) $$z_n = y_1 + \cdots + y_n \quad \forall n>1,$$

qualunque siano $n, K\in N$ risulta:

$$\|z_{n+k}-z_n\| = \|y_{n+1}+\dots+y_{n+k}\| \le \sum_{i=n+1}^{n+k} \|y_i\| < \sum_{i=n+1}^{n+k} 1/i^2 < \sum_{i=n+1}^{\infty} 1/i^2 .$$

Poiché la serie di termine generale $1/n^2$ é convergente, per ogni $\varepsilon > 0$ esiste $\nu \in N$ tale che:

$$\sum_{i=n+1}^{\infty} 1/i^2 < \varepsilon \qquad \forall n > \nu$$

e quindi risulta:

$$\|z_{n+k}-z_n\| < \varepsilon \qquad \forall n > \nu \text{ e } \forall k \in N .$$

Pertanto $(z_n)_{n\in N}$ é una successione di Cauchy di S, e quindi é convergente se S é uno spazio di Banach. Posto $z_0 = \lim_{n\to\infty} z_n$, sia X_0 l'elemento di S/H cui appartiene z_0.

Ricordando che $y_n \in X_{k_{n+1}} - X_{k_n}$, in virtu' della (1) si ha:

$$z_n \in X_{k_{n+1}} - X_{k_1} \qquad \forall n \in N$$

e quindi, tenendo presente l'ultima affermazione della (1.11) [pag.110]:

$$\lim_{n\to\infty} \left(X_{k_{n+1}} - X_{k_1} \right) = X_0$$

ossia:

$$\lim_{n\to\infty} X_{k_n} = X_{k_1} + X_0 .$$

Abbiamo così stabilito che esiste una successione estratta da $(X_n)_{n\in N}$ convergente: essendo $(X_n)_{n\in N}$ una successione di Cauchy, quanto sopra implica che anche $(X_n)_{n\in N}$ é convergente (1), e con ciò l'asserto é provato.

In virtù del criterio di convergenza di Cauchy in K^n, lo spazio pre-hilbertiano K^n è completo, e quindi è uno spazio di Hilbert.

Vedremo in seguito che ogni spazio vettoriale normato a dimensione finita è completo.

Si dimostra che gli spazi vettoriali normati considerati nel n.5 di pag.98 sono tutti completi, e quindi sono tutti spazi di Banach; più particolarmente, gli spazi:

(2) $\qquad l_2 \;,\; L^2(\Omega) \;,\; L^2(\Omega,\mu) \;,\; W^m(\Omega) \;,\; H^m(\Omega) \;,\; H_0^m(\Omega)$

sono spazi di Hilbert.

Anche $C^\infty(A)$, che è uno spazio vettoriale topologico metrizzabile, è completo.

Un esempio di spazio vettoriale non completo si ottiene associando allo spazio $C^m(\bar\Omega)$ [cfr. es.10 pag.96; Ω aperto di R^n, di misura finita], la norma definita da:

$$\|x\| = \left(\sum_{0\le|\gamma|\le m} \|D^\gamma x\|^p_{L^p(\Omega)} \right)^{1/p} \qquad (p \text{ reale} \ge 1),$$

ossia considerando ciascun $x \in C^m(\bar\Omega)$ come un elemento di $W^{m,p}(\Omega)$ (2) e ponendo $\|x\| = \|x\|_{W^{m,p}(\Omega)}$ [cfr. pag.103]. Lo spazio così ottenuto, che utilizzeremo in seguito per un esempio e che indicheremo

---

(1) Detta $(a_n)_{n\in N}$ una successione di punti di uno spazio metrico S, se tale successione è di Cauchy e se esiste una estratta $(a_{k_n})_{n\in N}$ convergente ad a_0, allora si ha $\lim a_n = a_0$. Basta infatti osservare che risulta, qualunque sia $n \in N$:

$$d(a_n, a_0) \le d(a_n, a_{k_n}) + d(a_{k_n}, a_0)$$

tenendo presente che, per ogni $\nu \in N$ se $n > \nu$ anche $k_n > \nu$ [in quanto $k_n \ge n \;\forall n \in N$].

(2) Ciò vuol dire che x si considera come una classe di funzioni, a due a due uguali quasi ovunque.

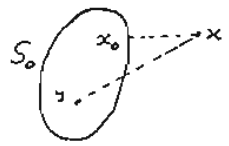
con $C^{m,p}(\Omega)$, non è completo perché si dimostra che $C^m(\bar\Omega)$ non è chiuso in $W^{m,p}(\Omega)$ [se lo spazio fosse completo, per il teor. (3.2) di pag. 112 dovrebbe essere $C^m(\bar\Omega)$ chiuso in $W^{m,p}(\Omega)$].
Ricordiamo che la chiusura di $C^m(\bar\Omega)$ in $W^{m,p}(\Omega)$ è lo spazio $H^{m,p}(\Omega)$ [cfr. es. 14 pag. 103].

4. Proiezione di un punto su un convesso chiuso in uno spazio di Hilbert. Proiezione ortogonale su un sottospazio chiuso. Proiettori.

Dimostriamo il seguente teorema:

(4.1) *Sia S_0 un sottoinsieme non vuoto, chiuso e convesso, di uno spazio di Hilbert S. Detta d la metrica per S generata dalla norma, per ogni $x \in S$ esiste uno ed un solo punto $x_0 \in S_0$ tale che:*

$$d(x,x_0) = d(x,S_0) = \inf_{y\in S_0} d(x,y)$$



ossia tale che:

$$\|x-x_0\| = \inf_{y\in S_0} \|x-y\| = \min_{y\in S_0} \|x-y\| . \tag{1}$$

Il punto x_0 si chiama *la proiezione di x su S_0*, e sarà denotato con

$$P_{S_0} x .$$

Dim. Posto:

$$\delta = \inf_{y\in S_0} \|x-y\| , \tag{2}$$

per ogni $n \in N$ esiste $y_n \in S_0$ tale che:

$$\delta \le \|x-y_n\| < \delta + \frac{1}{n} . \tag{3}$$

Dalla relazione [cfr. (4) pag. 97],

$$\|a-b\|^2 = 2\left(\|a\|^2+\|b\|^2\right) - \|a+b\|^2 \qquad \forall a,b \in S , \tag{4}$$

ponendo $a = y_n - x$, $b = y_m - x$ e tenendo presente che, siccome S_0 è convesso, il punto $\frac{1}{2}(y_n+y_m)$ appartiene ad S_0, si deduce, in base a (2) e (3):

$$\|y_n-y_m\|^2 = 2\left(\|y_n-x\|^2+\|y_m-x\|^2\right) - 4\left\|\tfrac{1}{2}(y_n+y_m)-x\right\|^2 <$$

$$< 2\left(\delta+\tfrac{1}{n}\right)^2 + 2\left(\delta+\tfrac{1}{m}\right)^2 - 4\delta^2 =$$

$$= 4\delta\left(\tfrac{1}{n}+\tfrac{1}{m}\right) + 2\left(\tfrac{1}{n^2}+\tfrac{1}{m^2}\right) .$$

Ne segue che $(y_n)_{n\in N}$ è una successione di Cauchy di S, e quindi è convergente; inoltre il suo limite, diciamolo x_0, appartiene ad S_0 in quanto S_0 è chiuso.

Tenendo presente (2) e (3) si ha allora:

$$\delta \le \|x-x_0\| \le \|x-y_n\|+\|y_n-x_0\| < \delta+\frac{1}{n}+\|y_n-x_0\| \qquad \forall n\in N$$

e quindi, passando al limite per $n\to\infty$:

$$\|x-x_0\| = \delta \quad ,$$

ossia la (1).

Allo scopo di far vedere che x_0 é l'unico punto di S_0 soddisfacente alla (1), supponiamo che la (1) stessa sussista anche con x_0' in luogo di x_0 (con $x_0' \in S_0$).

Ponendo in (4) $a = x_0 - x$, $b = x_0' - x$, si deduce allora:

$$\|x_0-x_0'\|^2 = 2\left(\|x_0-x\|^2+\|x_0'-x\|^2\right) - 4\left\|\tfrac{1}{2}(x_0+x_0')-x\right\|^2 =$$

$$= 4\delta^2 - 4\left\|\tfrac{1}{2}(x_0+x_0')-x\right\|^2 \le 0$$

e quindi $x_0 = x_0'$.

Supposto che S sia uno spazio di Hilbert reale, fissiamo un punto x^* di S e diciamo x_0 la proiezione di x^* sul sottoinsieme S_0 di S non vuoto, chiuso e convesso, sicché x_0 e' il punto in cui il funzionale:

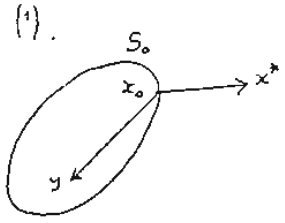
$$(5) \qquad F : y \in S_0 \longrightarrow \|x^*-y\|$$

assume il suo minimo valore, ossia, come suol dirsi, e' il punto che "minimizza" il funzionale (5). Essendo:

$$\|x^*-y\|^2 = (x^*-y, x^*-y) = \|x^*\|^2+\|y\|^2-2(x^*,y),$$

si può dire equivalentemente che x_0 e' l'unico punto di S_0 che minimizza il funzionale:

$$(6) \qquad \Phi : y \in S_0 \longrightarrow \tfrac{1}{2}\|y\|^2-(x^*,y) \qquad (1).$$



Dalla figura riportata qui accanto si desume poi che, nel caso $S=R^2$, se x_0 e' la proiezione di x^* su S_0, per ogni $y\in S_0$ i due vettori:

$$x^*-x_0 \quad , \quad y-x_0$$

formano un angolo ottuso, che può essere anche retto ma non acuto; di conseguenza il prodotto scalare dei due vettori e' ≤ 0, cioe' risulta:

$$(x^*-x_0, y-x_0) \le 0 \qquad \forall y\in S_0,$$

ossia:

$$(x_0, y-x_0) \ge (x^*, y-x_0) \qquad \forall y\in S_0 .$$

Il problema di determinare $x\in S_0$ tale da risultare:

$$(7) \qquad (x, y-x) \ge (x^*, y-x) \qquad \forall y\in S_0$$

si chiama una <u>disequazione variazionale</u>. Orbene, il fatto che x_0 e' la proiezione di x^* su S_0 equivale, non solo nel caso $S=R^2$ ma con riferimento ad uno spazio di Hilbert qualunque, al fatto che x_0 sia soluzione della (7). Precisamente sussiste il teorema che segue, nel quale riassumiamo anche l'osservazione sul funzionale (6):

---

(1) Minimizzare $\|x^*-y\|$ equivale a minimizzare $\|x^*-y\|^2$, quindi anche $\|y\|^2-2(x^*,y)$ che ne differisce per una costante, quindi anche $\frac{1}{2}\left[\|y\|^2-2(x^*,y)\right]$.

(4.2). Sia S_0 un sottoinsieme non vuoto, chiuso e convesso, di uno spazio di Hilbert reale S. Qualunque sia $x^* \in S$, le seguenti proposizioni sono equivalenti:

a) Il punto $x_0 \in S_0$ e' la proiezione di x^* su S_0.

b) Il punto $x_0 \in S_0$ e' l'unica soluzione della disequazione variazionale:

$$(x, y-x) \geq (x^*, y-x) \quad \forall y \in S_0 .$$

c) Il punto $x_0 \in S_0$ e' l'unico che minimizza il funzionale:

$$\Phi : y \in S_0 \to \frac{1}{2}\|y\|^2 - (x^*, y).$$

Dim. Si tratta di dimostrare che a) $\Leftrightarrow$ b), giacché l'equivalenza tra a) e c) e' stata gia' stabilita. Detta $\bar{b})$ la proposizione che si ottiene da b) sostituendo le parole "e' l'unica soluzione" con le parole "e' soluzione", proveremo che a) $\Leftrightarrow \bar{b})$ e da ciò, per l'unicità della proiezione, seguirà che a) $\Leftrightarrow$ b).

a) $\Rightarrow \bar{b})$. Fissato $y \in S_0$, per ogni $t \in [0,1]$ il punto $z(t) = (1-t)x_0 + ty$ appartiene ad S_0 in quanto S_0 e' convesso; essendo $z(0) = x_0$, la funzione reale:

$$\varphi(t) = \|z(t) - x^*\|^2$$

definita per $t \in [0,1]$ assume dunque per $t=0$ il suo minimo valore [in quanto per la a) il punto x_0 minimizza il funzionale (5)]. Allora deve risultare $\varphi'(0) \geq 0$, cioè, essendo:

$$\varphi'(t) = 2\,(z(t) - x^*, z'(t)) = 2\,(z(t) - x^*, y - x_0) \qquad (1),$$

deve risultare $(x_0 - x^*, y - x_0) \geq 0$. La $\bar{b})$ e' così dimostrata.

$\bar{b}) \Rightarrow$ a). Per la $\bar{b})$ risulta, qualunque sia $y \in S_0$:

$$(x^* - x_0, y - x_0) \leq 0$$

ossia:

$$(x^* - x_0, y - x^* + x^* - x_0) \leq 0$$

cioè:

$$(x^* - x_0, y - x^*) + \|x^* - x_0\|^2 \leq 0 .$$

Ne segue, per la disuguaglianza di Schwarz:

$$\|x^* - x_0\|^2 \leq (x^* - x_0, x^* - y) \leq \|x^* - x_0\| \cdot \|x^* - y\|$$

da cui, dividendo per $\|x^* - x_0\|$, si deduce:

$$\|x^* - x_0\| \leq \|x^* - y\| .$$

Siccome tale disuguaglianza sussiste per ogni $y \in S_0$, ne consegue la (1), ossia la a).

---

(1) Si verifica facilmente che, se $g(t)$ e' una funzione di una variabile reale a valori in uno spazio pre-hilbertiano S, ed e' derivabile, cioè tale che $\forall t$ esiste il limite:

$$\lim_{h \to 0} \frac{1}{h}\,[g(t+h) - g(t)] = g'(t),$$

la funzione reale di una variabile reale $\|g(t)\|^2$ e' derivabile, e risulta:

$$(*) \qquad D\|g(t)\|^2 = 2\,(g(t), g'(t)).$$

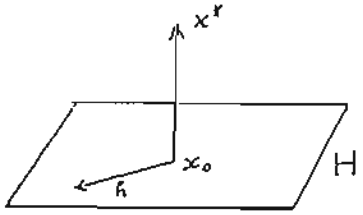
Nel nostro caso $g(t) = z(t) - x^*$, e risulta ovviamente $g'(t) = z'(t) = y - x_0$.

Per stabilire la (*) basta calcolare il limite del rapporto incrementale, tenendo presente che $\|g(t+h)\|^2 - \|g(t)\|^2 = (g(t+h), g(t+h)) - [g(t), g(t)]$. Sottrarre ed aggiungere $(g(t), g(t+h))$.

Dimostriamo ora il teorema seguente, che riguarda la proiezione di un punto su un sottospazio vettoriale chiuso di S (1):

(4.3) Se H e' un sottospazio vettoriale chiuso dello spazio di Hilbert S, qualunque sia $x^* \in S$ si ha, denotando con x_0 un punto di H:

x_0 = proiezione di x^* su H $\iff$ $x^* - x_0$ ortogonale ad H.



In altri termini, risulta:

(8) $\quad x^* - P_H x^*$ ortogonale ad H $\quad \forall x^* \in S$,

e $P_H x^*$ e' l'unico punto di H avente tale proprieta'.

Per tal motivo, $P_H x^*$ si chiama la proiezione ortogonale di x^* su H.

Notiamo che nel caso $S = R^3$ il teorema si interpreta geometricamente come nella figura.

Per semplicita', dimostriamo il teorema nel caso S reale.

Dim. Se x_0 e' la proiezione di x^* su H, per la b) del teor. (4.2) si ha:

(9) $\quad (x_0, y - x_0) \geq (x^*, y - x_0) \quad \forall y \in H$

da cui per $y = x_0 + h$ con $h \in H$ segue:

$$(x_0, h) \geq (x^*, h) \quad \forall h \in H.$$

Siccome tale relazione deve sussistere anche con $-h$ in luogo di h, ne segue che:

(10) $\quad (x_0, h) = (x^*, h) \quad \forall h \in H$,

ossia:

(10') $\quad (x^* - x_0, h) = 0 \quad \forall h \in H$.

Pertanto $x^* - x_0$ e' ortogonale ad H. Viceversa se vale la (10') con $x_0 \in H$, per ogni $y \in H$ ponendo in (10') $h = y - x_0$ si deduce che vale la (9), e quindi per il teor. (4.2) x_0 e' la proiezione di x^* su H.

— — — — — — — —

Si osservi che, essendo la (9) equivalente alla (10'), per il teor. (4.3) la proiezione ortogonale x_0 di x^* su H e' l'unica soluzione dell'equazione nell'incognita $x \in H$:

$$(x, h) = (x^*, h) \quad \forall h \in H$$

che prende il nome di equazione variazionale.

Nel caso che H sia il sottospazio generato da un insieme ortonormale finito $\{e_1, e_2, \dots, e_n\}$ di vettori di S (2), H ha dimensione finita, quindi e' chiuso [cfr. (2.7) pag. 87], e per ogni $x \in S$ la proiezione ortogonale $P_H x$ di x su H si esprime come combinazione lineare dei vettori $e_1, e_2, \dots, e_n$ mediante il prodotto scalare in S. Precisamente, dimostriamo che risulta:

(11) $\quad P_H x = (x, e_1) e_1 + (x, e_2) e_2 + \cdots + (x, e_n) e_n \quad \forall x \in S.$

In virtu' della (3) di pag. 96 (3), dalla (11) consegue poi:

(12) $\quad \|P_H x\|^2 = |(x, e_1)|^2 + |(x, e_2)|^2 + \cdots + |(x, e_n)|^2 \quad \forall x \in S.$

---

(1) Ovviamente un sottospazio di S e' un insieme (non vuoto e) convesso.

(2) L'insieme $\{e_1, \dots, e_n\}$ si dice ortonormale quando e' costituito da vettori a due a due ortogonali aventi norma uguale ad 1. Si veda la nota (1) a piè di pag. 107.

(3) E' facile verificare che la citata (3), stabilita con riferimento a due vettori ortogonali di uno spazio pre-hilbertiano, si estende ad n vettori a due a due ortogonali (naturalmente nell'ipotesi che esistano): basta procedere induttivamente.

Dim. (11). Siccome H è il sottospazio generato da $\{e_1,\dots,e_n\}$ e $P_H x$ appartiene ad H, esistono n scalari $c_1,\dots,c_n$ tali da risultare:

(13) $$P_H x = c_1 e_1 + c_2 e_2 + \cdots + c_n e_n .$$

Per ogni $i \in \{1,\dots,n\}$, moltiplicando ambo i membri scalarmente per e_i e ricordando che i vettori $e_1,\dots,e_n$ sono a due a due ortogonali, si ha:

(14) $$(P_H x, e_i) = c_i (e_i, e_i) = c_i \|e_i\|^2 = c_i .$$

D'altra parte il vettore $x - P_H x$ è ortogonale ad H [teor. (4.3)], ed e_i è un vettore di H, sicché risulta:

$$(x - P_H x, e_i) = 0$$

e quindi:

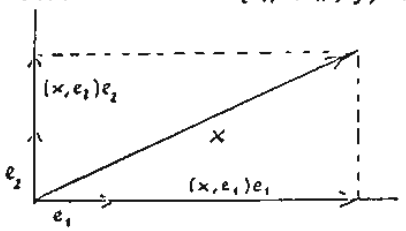
$$(P_H x, e_i) = (x, e_i).$$

Dalla (14) consegue allora $c_i = (x, e_i)$ e ciò, per la (13), prova la (11).

Notiamo che se $x \in H$ [sottospazio generato dall'insieme ortonormale $\{e_1,\dots,e_n\}$], e solo in tal caso, la (11) diviene:

(15) $$x = (x,e_1)e_1 + (x,e_2)e_2 + \cdots + (x,e_n)e_n .$$

I numeri $(x,e_1), (x,e_2), \dots, (x,e_n)$ si chiamano perciò <u>le componenti del vettore x secondo i vettori $e_1, e_2, \dots, e_n$</u>.



Nel caso $H = R^2$ la (15) si interpreta geometricamente come nella figura.

Tornando a considerare un sottospazio vettoriale chiuso H dello spazio di Hilbert S, osserviamo che, siccome il complemento ortogonale $H^\perp$ è un sottospazio vettoriale chiuso di S [indipendentemente dal fatto che H è un sottospazio, e che è chiuso, cfr. (1.9) pag. 109], per ogni $x \in S$ ha senso considerare sia la proiezione ortogonale di x su H, sia la proiezione ortogonale di x su $H^\perp$. Orbene, si verifica subito che:

(4.4) <u>Se H è un sottospazio vettoriale chiuso dello spazio di Hilbert S, risulta:</u>

(16) $$x = P_H x + P_{H^\perp} x \qquad \forall x \in S ,$$

<u>e la (16) è l'unica decomposizione di x nella somma di un vettore di H e di un vettore di $H^\perp$.</u>

Dim. Fissato $x \in S$ e posto $x_0 = x - P_H x$ [sicché per la (4.3) x_0 è ortogonale ad H, ossia è un punto di $H^\perp$], per provare la (16) si tratta di far vedere che $x_0 = P_{H^\perp} x$, cioè, per la (4.3) con $H^\perp$ in luogo di H e $x^* = x$, che il vettore $x - x_0$ è ortogonale ad $H^\perp$:

$$(x - x_0, h') = 0 \quad \forall h' \in H^\perp .$$

Ma ciò è ovvio, perché ciascun elemento h' di $H^\perp$ è ortogonale a tutti gli elementi di H (per la definizione di $H^\perp$) e $x - x_0$ è un elemento di H [in quanto uguale a $P_H x$).

Quanto all'unicità della decomposizione (16), basta osservare che se risulta:

$$x = h + h' \quad \text{con } h \in H, h' \in H^\perp ,$$

il vettore $x - h$ appartiene ad $H^\perp$ e quindi è ortogonale a tutti gli elementi di H. Essendo $x - h$ ortogonale ad H e $h \in H$, per il teor. (4.3) si ha $h = P_H x$, e per la (16) si ha poi $h' = P_{H^\perp} x$.

Dunque se H è un sottospazio vettoriale chiuso di S, lo spazio S è somma diretta dei sottospazi H ed $H^\perp$; in altri termini, tali sottospazi sono supplementari. Perciò il complemento ortogonale $H^\perp$ di un sottospazio vettoriale chiuso H si chiama anche il supplemento ortogonale di H.

Notiamo che dalla (16) consegue [cfr. (3) pag. 96]:

$$\|x\|^2 = \|P_H x\|^2 + \|P_{H^\perp} x\|^2 \qquad \forall x \in S \tag{17}$$

e che tale relazione implica:

$$\|P_H x\| \leq \|x\| \qquad \forall x \in S, \tag{18}$$

valendo l'uguaglianza se e solo se $x \in H$ [se vale l'uguaglianza, per la (17) si ha $P_{H^\perp} x = 0$ e quindi, per la (16), $x = P_H x$, da cui $x \in H$].

Dalla (18) si deduce poi una disuguaglianza relativa al caso che H sia il sottospazio generato da un insieme ortonormale finito $\{e_1, \dots, e_n\}$. Precisamente, in virtù della (18), la (12) di pag. 118 fornisce:

$$\sum_{i=1}^{n} |(x, e_i)|^2 \leq \|x\|^2 \qquad \forall x \in S, \tag{19}$$

e nella (19) vale l'uguaglianza se e solo se $x \in H$ [per quanto osservato sulla (18)].

Qualunque sia il sottospazio vettoriale H di S, non necessariamente chiuso, il complemento ortogonale $H^\perp$ è un sottospazio vettoriale chiuso [cfr. (1.9) pag. 109], e quindi il suo complemento ortogonale può chiamarsi il supplemento ortogonale di $H^\perp$. Orbene, si ha che:

(4.5) Per ogni sottospazio vettoriale H dello spazio di Hilbert S, il supplemento ortogonale di $H^\perp$ è la chiusura di H; in altri termini, i sottospazi:

$$H^\perp \ , \quad \overline{H}$$

sono supplementari.

Dim. Poichè la chiusura di H è un sottospazio vettoriale chiuso di S, sappiamo che i due sottospazi:

$$\overline{H} \ , \quad \overline{H}^\perp$$

sono supplementari. Si tratta allora di far vedere che:

$$H^\perp = \overline{H}^\perp,$$

ossia che ogni vettore ortogonale ad H è anche ortogonale alla chiusura di H (essendo banale il viceversa).

Fissato x ortogonale ad H, sia h un arbitrario vettore di $\overline{H}$. Per la (1.7) di pag. 108 esiste allora una successione $(h_n)_{n \in \mathbb{N}}$ di vettori di H convergente ad h, e dalla relazione $(x, h_n) = 0 \ \forall n \in \mathbb{N}$ consegue $(x, h) = 0$ [cfr. (5) pag. 107]. Dunque x è anche ortogonale alla chiusura di H, e ciò prova l'asserto.

Rileviamo una importante conseguenza della (4.5):

(4.6). Per ogni sottospazio vettoriale H dello spazio di Hilbert S, sussiste l'equivalenza:

$$\overline{H} = S \iff H^{\perp} = \{0\}.$$

In altri termini, se non esiste alcun vettore diverso dal vettore nullo che sia ortogonale ad H, allora H è denso in S.

Dim. Poichè ovviamente i sottospazi S e $\{0\}$ sono l'uno il supplementare ortogonale dell'altro, se il primo di essi coincide con la chiusura di H, per il teor. (4.5) l'altro coincide con $H^{\perp}$, e se il secondo coincide con $H^{\perp}$, per lo stesso teorema l'altro coincide con $\overline{H}$.

Abbiamo fin qui implicitamente denotato con P_H, per ogni sottospazio vettoriale chiuso H dello spazio di Hilbert S, l'operatore di S in H:

$$x \in S \to (\text{la proiezione ortogonale di } x \text{ su } H).$$

Tale operatore si chiama un proiettore (precisamente, il proiettore di S su H).

Osserviamo che:

(4.7) Il proiettore P_H di S su H è un operatore lineare.

Dim. Qualunque siano $x_1, x_2 \in S$, dalle relazioni [cfr. (16) pag. 119]:

$$x_1 = P_H x_1 + P_{H^{\perp}} x_1 \quad , \quad x_2 = P_H x_2 + P_{H^{\perp}} x_2$$

consegue:

$$x_1 + x_2 = \left[P_H x_1 + P_H x_2 \right] + \left[P_{H^{\perp}} x_1 + P_{H^{\perp}} x_2 \right].$$

Tale uguaglianza realizza una decomposizione del vettore $x_1 + x_2$ nella somma di un vettore di H e di un vettore di $H^{\perp}$. Poichè l'unica decomposizione di questo tipo è la (16) di pag. 119 con $x = x_1 + x_2$, ne segue che:

$$P_H x_1 + P_H x_2 = P_H (x_1 + x_2).$$

D'altra parte, qualunque siano $x \in S$ e $\lambda \in K$, per la citata (16) si ha:

$$\lambda x = \lambda P_H x + \lambda P_{H^{\perp}} x \ ,$$

e con lo stesso ragionamento fatto dianzi si deduce che $\lambda P_H x$ è la proiezione ortogonale su H del vettore a primo membro, cioè che:

$$\lambda P_H x = P_H (\lambda x).$$

Concludiamo questo numero con un esempio di proiezione ortogonale.

Sia $S = L^2(]0,1[)$, e sia H il sottospazio vettoriale di S costituito dalle restrizioni a $]0,1[$ dei polinomi

nel campo reale di grado ≤ 1. Il sottospazio H ha dimensione 2, quindi è chiuso [cfr. (2.7) pag. 87].

Posto $x^*(t) = t^2 \ \forall t \in]0,1[$, determiniamo la proiezione ortogonale del vettore x^* di S sul sottospazio H: diciamola x_0, e poniamo $x_0(t) = at+b \ \forall t \in]0,1[$.

Siccome x_0 è il punto di H tale che $x^* - x_0$ sia ortogonale a tutti gli elementi di H, detto h un arbitrario elemento di H e posto $h(t) = \lambda t + \mu \ \forall t \in]0,1[$ deve allora risultare, ricordando il prodotto scalare in $L^2(]0,1[)$:

$$\int_0^1 (t^2 - at - b)(\lambda t + \mu)\, dt = 0 \quad \forall \lambda, \mu \in \mathbb{R} .$$

Con semplici calcoli se ne deduce che $a = 1$, $b = -\frac{1}{6}$, e quindi la proiezione ortogonale di x^* su H è il polinomio $x_0(t) = t - \frac{1}{6}$.

5. Sistemi ortonormali di punti di uno spazio pre-hilbertiano.

Sia S uno spazio pre-hilbertiano. Se $(x_i)_{i \in I}$ è una famiglia di vettori di S tale che x_r ed x_s sono ortogonali per $r \neq s$, si dice che $(x_i)_{i \in I}$ è una <u>famiglia ortogonale</u>, oppure un <u>sistema ortogonale</u>; l'insieme dei vettori della famiglia si dice poi un <u>insieme ortogonale</u> di vettori di S.

Una famiglia di vettori $(e_i)_{i \in I}$ si dice una <u>famiglia ortonormale</u>, oppure un <u>sistema ortonormale</u>, quando è ortogonale ed inoltre tutti i vettori della famiglia hanno norma uguale ad 1. L'insieme dei vettori della famiglia si dice poi un <u>insieme ortonormale</u>.

Osserviamo che:

(5.1). <u>I vettori di ogni sistema ortonormale sono linearmente indipendenti.</u>

— — — — — — — — —

Dim. Sia $(e_i)_{i \in I}$ un sistema ortonormale. Si tratta di far vedere che per ogni sottoinsieme finito di I, sia esso $\{i_1, \dots, i_n\}$, i vettori $e_{i_1}, \dots, e_{i_n}$ sono linearmente indipendenti.

Supponiamo dunque che $\lambda_1, \dots, \lambda_n$ siano scalari tali che:

$$\lambda_1 e_{i_1} + \lambda_2 e_{i_2} + \cdots + \lambda_n e_{i_n} = 0 . \tag{1}$$

Qualunque sia $s \in \{1, \dots, n\}$, moltiplicando scalarmente per e_{i_s} ambo i membri della (1) si ottiene:

$$\lambda_s (e_{i_s}, e_{i_s}) = 0 ,$$

in quanto per $r \neq s$ risulta $i_r \neq i_s$ e quindi $(e_{i_r}, e_{i_s}) = 0$. Essendo poi $(e_{i_s}, e_{i_s}) = \|e_{i_s}\|^2 = 1$, ne segue $\lambda_s = 0$ e ciò, essendo s arbitrario, prova l'asserto.

— — — — — — — — —

Per quanto attiene all'esistenza di sistemi ortonormali in uno spazio pre-hilbertiano sussiste il teorema seguente, la cui dimostrazione contiene un metodo per costruire una famiglia ortonormale a partire da una famiglia finita o numerabile di vettori linearmente indipendenti (<u>procedimento di ortonormalizzazione di Schmidt</u>):

(5.2). In ogni spazio di Hilbert S esistono insiemi ortonormali.
Se S ha dimensione finita, ogni insieme ortonormale è finito.
Se S ha dimensione infinita, esistono insiemi ortonormali finiti ed insiemi ortonormali numerabili; nell'ulteriore ipotesi che lo spazio sia separabile, non esistono insiemi ortonormali aventi potenza maggiore del numerabile.

— — — — — — — — —

Dim. Supponiamo dapprima che S abbia dimensione infinita: esiste allora una successione $(x_n)_{n\in\mathbb{N}}$ di vettori di S linearmente indipendenti.

Poniamo:

$$y_1 = x_1 \qquad e_1 = \frac{y_1}{\|y_1\|} \qquad H_1 = \text{sottospazio generato da } \{e_1\}$$

$$y_2 = x_2 - P_{H_1} x_2 \qquad e_2 = \frac{y_2}{\|y_2\|} \qquad H_2 = \text{sottospazio generato da } \{e_1, e_2\}$$

$$y_3 = x_3 - P_{H_2} x_3 \qquad e_3 = \frac{y_3}{\|y_3\|} \qquad H_3 = \text{sottospazio generato da } \{e_1, e_2, e_3\}$$

. . . . . .

$$y_n = x_n - P_{H_{n-1}} x_n \qquad e_n = \frac{y_n}{\|y_n\|} \qquad H_n = \text{sottospazio generato da } \{e_1, e_2, \dots, e_n\}$$

. . . . . .

Le uguaglianze che definiscono induttivamente $e_1, e_2, e_3 \dots$ hanno senso in quanto i vettori $y_1, y_2, y_3 \dots$ sono tutti diversi da 0: infatti si verifica induttivamente che ciascun vettore y_n è una combinazione lineare dei vettori $x_1, x_2, \dots, x_n$ con il coefficiente di x_n uguale ad 1 [perciò se fosse $y_n = 0$ sarebbe x_n combinazione lineare di $x_1, \dots, x_{n-1}$, il che contraddice il fatto che i vettori della successione $(x_n)_{n\in\mathbb{N}}$ sono linearmente indipendenti].

Proviamo che $(e_n)_{n\in\mathbb{N}}$ è una successione ortonormale.

Essendo ovvio che ciascun vettore e_n ha norma uguale ad 1, basta far vedere che ciascun vettore e_n è ortogonale a tutti i precedenti, cioè che:

$$(e_n, e_k) = 0 \quad \forall k \in \{1, 2, \dots, n-1\} \qquad (n > 1).$$

Si ha:

$$(e_n, e_k) = \frac{1}{\|y_n\|} (y_n, e_k)$$

e d'altra parte $(y_n, e_k) = 0$ $\forall k \in \{1, \dots, n-1\}$ perché y_n è ortogonale al sottospazio H_{n-1} [cfr. (8) pag. 118] e il vettore e_k appartiene a tale sottospazio.

Siccome la successione $(e_n)_{n\in\mathbb{N}}$ è invertibile (in quanto ortonormale), resta acquisito che in S esistono insiemi ortonormali numerabili.

Il precedente procedimento è ancora valido se si sostituisce la successione $(x_n)_{n\in\mathbb{N}}$ con una famiglia finita di vettori linearmente indipendenti: pertanto in S esistono anche insiemi ortonormali finiti; se S ha dimensione finita, vale la stessa conclusione, e per il Teor. (5.1) non esistono insiemi ortonormali infiniti.

Resta dunque da dimostrare l'ultima affermazione dell'enunciato, e perciò supponiamo che S sia separabile. Denotiamo quindi con E un qualunque insieme ortonormale di vettori di S.

Osserviamo che per $x', x'' \in E$, con $x' \neq x''$, risulta:

$$\|x'-x''\|^2 = (x'-x'', x'-x'') = \|x'\|^2 - (x', x'') - (x'', x') + \|x''\|^2 = 2$$

e quindi:

$$\|x'-x''\| = \sqrt{2} \quad . \tag{2}$$

Detto Y un insieme numerabile denso in S, per ogni $x \in E$ denotiamo con y_x un elemento di Y tale che:

$$\|x - y_x\| < \frac{\sqrt{2}}{3}$$

e consideriamo l'insieme degli elementi della famiglia $(y_x)_{x \in E}$, diciamolo Y_0.

L'applicazione:

$$f : x \in E \to y_x \in Y_0$$

è biunivoca. Infatti qualunque siano $x', x'' \in E$, con $x' \neq x''$, ricordando la (2) si ha:

$$\sqrt{2} = \|x'-x''\| \leq \|x'-y_{x'}\| + \|y_{x'}-y_{x''}\| + \|y_{x''}-x''\| < \frac{\sqrt{2}}{3} + \|f(x')-f(x'')\| + \frac{\sqrt{2}}{3}$$

e quindi:

$$\|f(x')-f(x'')\| > \frac{\sqrt{2}}{3} > 0 .$$

Dunque gli insiemi E, Y_0 sono equipotenti. Essendo Y_0 finito o numerabile in quanto contenuto nell'insieme numerabile Y, l'asserto è dimostrato.

— — — — — — — — — —

Notiamo che la successione $(y_n)_{n \in N}$ costruita nella dimostrazione del teorema (5.2) è una successione ortogonale. La successione $(e_n)_{n \in N}$ costruita a partire dalla $(y_n)_{n \in N}$ si dice ottenuta <u>normalizzando</u> la successione ortogonale $(y_n)_{n \in N}$.

Un sistema ortonormale $(e_i)_{i \in I}$ di vettori di S si dice <u>completo</u> quando non esiste alcun vettore di $S-\{0\}$ che sia ortogonale a ciascun elemento della famiglia $(e_i)_{i \in I}$.

Dal teor. (4.6) di pag. 121 segue che:

(5.3) <u>Un sistema ortonormale $(e_i)_{i \in I}$ di vettori di uno spazio di Hilbert S è completo se e solo se il sottospazio da essi generato è denso in S.</u>

Se lo spazio di Hilbert S ha dimensione finita, esiste ovviamente in S un sistema ortonormale completo (1). Per gli spazi a dimensione infinita sussiste il seguente teorema, che ci limitiamo ad enunciare:

(5.4) <u>Se S è uno spazio di Hilbert di dimensione infinita, l'esistenza di una successione di vettori di S che sia un sistema ortonormale completo equivale al fatto che lo spazio sia separabile.</u>

---

(1) Basta considerare una base di S, e applicare il procedimento di ortonormalizzazione di Schmidt. Se n è la dimensione di S, si ottiene un insieme ortonormale di n vettori: il sottospazio da essi generato è denso in S perché coincide con S.

6. Serie in uno spazio di Banach. Serie di Fourier in uno spazio di Hilbert.

Sia $(x_n)_{n\in N}$ una successione di punti di uno spazio vettoriale topologico di Hausdorff S.
Si dice che la serie :

$$(1) \qquad x_1 + x_2 + \cdots + x_n + \cdots$$

è convergente , quando la successione delle somme parziali :

$$\left(\sum_{i=1}^{n} x_i \right)_{n\in N}$$

è convergente. In tale ipotesi, il limite:

$$\lim_{n\to\infty} \sum_{i=1}^{n} x_i$$

si chiama somma della serie (1) , e si denota con

$$\sum_{i=1}^{\infty} x_i \ .$$

Osserviamo che:

(6.1) Nell'ipotesi che S sia uno spazio di Banach, la serie (1) è convergente se e solo se è verificata la seguente condizione:

Per ogni $\varepsilon>0$ esiste $\nu\in N$ tale che:

$$\| x_{n+1} + x_{n+2} + \cdots + x_{n+k} \| < \varepsilon \quad \forall n>\nu \ e \ \forall k \in N .$$

Dim. Basta osservare che, detta $(s_n)_{n\in N}$ la successione delle somme parziali della (1), la condizione dell'enunciato equivale al fatto che $(s_n)_{n\in N}$ è una successione di Cauchy di S .

Nelle considerazioni che seguono riterremo sottinteso che lo spazio S non sia a dimensione finita.

(6.2) Sia S uno spazio di Hilbert, e sia $(h_n)_{n\in N}$ una successione ortogonale di vettori di S .

La serie :

$$h_1 + h_2 + \cdots + h_n + \cdots$$

è convergente se e solo se la successione numerica $(\|h_n\|)_{n\in N}$ è un elemento di l_2 [1] ; in altri termini, se e solo se la serie numerica :

$$\|h_1\|^2 + \|h_2\|^2 + \cdots + \|h_n\|^2 + \cdots$$

è convergente.

---

(1) Cfr. 2) pag. 99.

Dim. Essendo $(h_n)_{n\in N}$ una successione ortogonale, in virtù della (3) a pag. 96 (1) qualunque siano $n, k \in N$ risulta:

$$\| h_{n+1} + h_{n+2} + \cdots + h_{n+k} \|^2 = \| h_{n+1} \|^2 + \| h_{n+2} \|^2 + \cdots + \| h_{n+k} \|^2 .$$

Dopo ciò, l'asserto consegue dalla (6.1) e dal criterio di Cauchy per le serie numeriche.

(6.3) Siano : S uno spazio di Hilbert, $(e_n)_{n\in N}$ una successione ortonormale di vettori di S.

Detta $(c_n)_{n\in N}$ una successione di scalari, la serie:

$$(2) \qquad c_1 e_1 + c_2 e_2 + \cdots + c_n e_n + \cdots$$

è convergente se e solo se $(c_n)_{n\in N}$ appartiene ad l_2.

Se la (2) è convergente, detta x la sua somma si ha :

$$(3) \qquad (x, e_n) = c_n \qquad \forall n \in N .$$

Dim. Poiché $(c_n e_n)_{n\in N}$ è una successione ortogonale, in virtù della (6.2) la serie (2) è convergente se e solo se la successione numerica:

$$(4) \qquad (\| c_n e_n \|)_{n\in N}$$

appartiene ad l_2. Essendo $\| e_n \| = 1 \quad \forall n \in N$, in quanto $(e_n)_{n\in N}$ è ortonormale, la (4) è la successione $(|c_n|)_{n\in N}$, che appartiene ad l_2 se e solo se $(c_n)_{n\in N}$ appartiene ad l_2.

Supponiamo la (2) convergente, e denotiamo con x la sua somma.

Per ogni $k \in N$ e per ogni $n \geq k$ si ha:

$$\left(\sum_{i=1}^{n} c_i e_i , e_k \right) = \sum_{i=1}^{n} c_i (e_i, e_k) = c_k$$

e quindi, passando al limite per $n \to \infty$ e ricordando la (5) a pag. 107 :

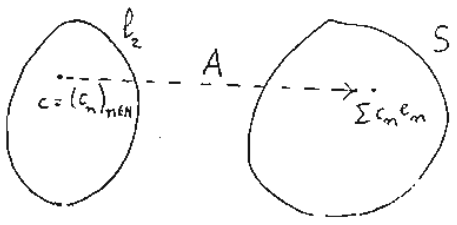
$$(x, e_k) = c_k .$$

In virtù del teor. (6.3) ha senso considerare la corrispondenza A che ad ogni punto di l_2 :

$$c = (c_n)_{n\in N}$$

associa la somma della serie (2), che è un elemento dello spazio di Hilbert S :

$$A : c = (c_n)_{n\in N} \in l_2 \longrightarrow \sum_{n=1}^{\infty} c_n e_n \in S .$$



---

(1) Cfr. anche la nota (3) a pag. 118.

Ovviamente A è un operatore lineare di l_2 in S. L'operatore A è anche invertibile [cfr. (2') pag.73]: infatti se $c=(c_n)_{n\in N}$ è un punto di l_2 tale che $Ac=0$, ossia tale che $\sum_{n=1}^{\infty} c_n e_n = 0$, per il teor. (6.3) [cfr. (3)] risulta $c_n=(0,e_n)=0$, ossia $c=0$.

E' naturale allora domandarsi se A è anche suriettivo, cioè se per ogni $x\in S$ esiste $c=(c_n)_{n\in N}\in l_2$ tale da risultare:

$$Ac = x \,. \tag{5}$$

L'eventuale risposta affermativa comporterebbe che ogni vettore x di S si possa esprimere nella forma:

$$x = c_1 e_1 + c_2 e_2 + \cdots + c_n e_n + \cdots \,, \tag{6}$$

come accade se per almeno un $n\in N$ il vettore x appartiene al sottospazio H_n di S generato da $\{e_1, e_2, \ldots, e_n\}$ [1], nel qual caso la serie si riduce alla sua somma parziale n^{ma}. Gli scalari $c_1, c_2, \ldots, c_n, \ldots$ si potrebbero allora chiamare "le componenti del vettore x", oppure "le coordinate del punto x", e gli elementi dello spazio S si presenterebbero come vettori con un'infinità numerabile di componenti, o come punti con un'infinità numerabile di coordinate [esattamente come accade per l_2, che però è uno spazio particolare].

Orbene, affinché esista $c\in l_2$ tale da verificare la (5), per il teor. (6.3) [cfr. (3)] è necessario che l'elemento c di l_2 sia il seguente:

$$c = \big((x,e_n)\big)_{n\in N} \,. \tag{7}$$

Ma la (7) è sufficiente per il verificarsi della (5)? In altri termini, ci domandiamo: Si verificano le seguenti circostanze?

a) La successione a secondo membro della (7) è un elemento di l_2.

b) La serie, convergente in virtù della a):

$$(x,e_1)e_1 + (x,e_2)e_2 + \cdots + (x,e_n)e_n + \cdots \,, \tag{8}$$

ha per somma x.

Per ogni $x\in S$ la (8) si chiama <u>la serie di Fourier di x rispetto al sistema ortonormale</u> $(e_n)_{n\in N}$, e i termini della successione numerica (7) si chiamano <u>i coefficienti di Fourier di x rispetto al sistema ortonormale</u> $(e_n)_{n\in N}$.

<u>La circostanza a) si verifica qualunque sia $x\in S$</u>. Infatti per ogni $n\in N$ la somma parziale n^{ma} della serie:

$$|(x,e_1)|^2 + |(x,e_2)|^2 + \cdots + |(x,e_n)|^2 + \cdots \tag{9}$$

è maggiorata da $\|x\|^2$, in virtù della (19) di pag.120, e perciò la serie, a termini non negativi, è convergente [il che significa, appunto, che la successione a secondo membro della (7) è un elemento di l_2]. Per la somma della serie (9) vale poi la maggiorazione:

---

[1] Cfr. (15) pag.119.

$$(10) \qquad \sum_{n=1}^{\infty} |(x,e_n)|^2 \le \|x\|^2 \qquad \forall x \in S$$

che prende il nome di disuguaglianza di Bessel.

Invece la circostanza b) non si verifica qualunque sia $x \in S$; cioè la somma della serie di Fourier di x rispetto al sistema ortonormale $(e_n)_{n\in N}$ non coincide con x per qualunque $x \in S$, se si fa riferimento ad un sistema ortonormale qualsiasi.

Vedremo tra breve che la b) è vera per ogni $x \in S$ se e solo se il sistema ortonormale $(e_n)_{n\in N}$ è completo [secondo la definizione introdotta a pag. 124]. Intanto osserviamo quanto segue.

Detto H_n, per ogni $n \in N$, il sottospazio di S generato da $\{e_1, \dots, e_n\}$, per la (11) di pag. 118 risulta:

$$(11) \qquad P_{H_n} x = (x,e_1)e_1 + (x,e_2)e_2 + \cdots + (x,e_n)e_n \qquad \forall x \in S,$$

sicchè per ciascun $x \in S$ la somma parziale n^{ma} della serie di Fourier di x è la proiezione ortogonale di x su H_n.

Sussiste poi il teorema:

(6.4) Per ciascun punto x dello spazio di Hilbert S, la somma della serie di Fourier di x rispetto al sistema ortonormale (e_n) è la proiezione ortogonale di x sulla chiusura del sottospazio H generato da $(e_n)_{n\in N}$. Risulta cioè:

$$(12) \qquad \lim_{n\to\infty} P_{H_n} x = P_{\overline{H}} x \qquad \forall x \in S.$$

La somma della serie coincide con x se e solo se risulta:

$$(13) \qquad \sum_{n=1}^{\infty} |(x,e_n)|^2 = \|x\|^2 \qquad \text{(uguaglianza di Parseval)} \ (1).$$

Dim. In virtù della (8) di pag. 118, qualunque sia $n \in N$ il vettore $x - P_{H_n} x$ è ortogonale ad H_n, e quindi è ortogonale al vettore e_k se $k \le n$. Si ha allora:

$$(14) \qquad (x - P_{H_n} x \,,\, e_k) = 0 \qquad \forall n \ge k.$$

Detta y la somma della serie (8), ossia posto $y = \lim_{n\to\infty} P_{H_n} x$, dalla (14) passando al limite per $n \to \infty$ si deduce [cfr. (5) pag. 107]:

$$(x - y, e_k) = 0.$$

Poichè tale relazione sussiste qualunque sia $k \in N$, ne segue che il vettore $x-y$ è ortogonale ad H. D'altra parte ogni elemento h di $\overline{H}$ è il limite di una successione $(h_n)_{n\in N}$ di elementi di H, e da $(x-y, h_n) = 0 \ \forall n \in N$ segue $(x-y, h) = 0$: pertanto $x-y$ è ortogonale alla chiusura di H, e ciò vuol dire che y è la proiezione ortogonale di x su $\overline{H}$ [cfr. (4.3) pag. 118 con $\overline{H}$ in luogo di H e $x^* = x$, $x_0 = y$].

---

(1) In altri termini, se e solo se, per il considerato punto x di S, nella disuguaglianza (10) di Bessel vale il segno di uguaglianza.

Per dimostrare l'ultima affermazione dell'enunciato, denotiamo con s_n la somma parziale n^{ma} della serie (8). Dalla relazione:

$$x = (x - s_n) + s_n$$

consegue (1):

$$\|x\|^2 = \|x - s_n\|^2 + \|s_n\|^2$$

ossia, essendo $s_n = P_{H_n} x$ e ricordando la (12) di pag. 118:

$$\|x\|^2 = \|x - s_n\|^2 + \sum_{i=1}^{n} |(x, e_i)|^2 .$$

Da tale relazione si deduce l'asserto.

— — — — — — — —

Aggiungiamo infine l'osservazione seguente.

Fissati $x \in S$ ed $n \in N$, consideriamo la funzione reale φ definita in K^n dalla relazione:

$$\varphi(c_1, c_2, \dots, c_n) = \|x - (c_1 e_1 + c_2 e_2 + \cdots + c_n e_n)\| \qquad \forall (c_1, \dots, c_n) \in K^n .$$

Siccome al variare di $(c_1, \dots, c_n)$ in K^n il punto $y = c_1 e_1 + \cdots + c_n e_n$ descrive il sottospazio H_n, essendo per definizione di proiezione:

$$\|x - P_{H_n} x\| = \min_{y \in H_n} \|x - y\| = \min_{K^n} \varphi ,$$

si può dire che: per ogni $x \in S$ e per ogni $n \in N$ la n^{pla} ordinata dei primi n coefficienti di Fourier di x minimizza in tutto K^n la funzione $\varphi(c_1, c_2, \dots, c_n)$.

Dal teorema (6.4) si deduce subito, con riferimento alla circostanza b) considerata a pag. 127, la conclusione anticipata a pag. 128. Siccome si ha, ovviamente:

$$x = P_{\overline{H}} x \quad \forall x \in S \iff x \in \overline{H} \quad \forall x \in S \iff \overline{H} = S ,$$

dalla (12) segue che, se e solo se $\overline{H} = S$, cioè se e solo se il sistema ortonormale $(e_n)_{n \in N}$ è completo [cfr. (5.3) pag. 124], la serie di Fourier di x rispetto ad $(e_n)_{n \in N}$ ha per somma x qualunque sia $x \in S$.

Riassumiamo tale conclusione, insieme ad una ovvia conseguenza dell'ultima affermazione della (6.4), col seguente enunciato:

(6.5) Sia $(e_n)_{n \in N}$ un sistema ortonormale di vettori dello spazio di Hilbert S. Le seguenti proposizioni sono equivalenti:

1) La serie di Fourier di x rispetto ad $(e_n)_{n \in N}$ ha per somma x qualunque sia $x \in S$.

2) Il sottospazio generato da $(e_n)_{n \in N}$ è denso in S.

3) L'uguaglianza di Parseval relativa ad $(e_n)_{n \in N}$ sussiste per ogni $x \in S$.

4) Il sistema ortonormale $(e_n)_{n \in N}$ è completo.

---

(1) I vettori $x - s_n$ ed s_n sono ortogonali: infatti s_n appartiene al sottospazio H_n, e per la (11) si ha $x - s_n = x - P_{H_n} x$, che è ortogonale ad H_n [cfr. (8) pag. 118].

Dopo ciò, si ricordi la (3) di pag. 96.

Ricordiamo che, in uno spazio di Hilbert S a dimensione infinita, un sistema ortonormale numerabile $(e_n)_{n\in N}$ che sia completo esiste se e solo se lo spazio S è separabile [cfr. (5.4) pag. 124]. In tale ipotesi, fissato un sistema ortonormale completo $(e_n)_{n\in N}$, consideriamo la corrispondenza F che ad ogni $x\in S$ associa la successione $(x,e_n)_{n\in N}$ dei coefficienti di Fourier di x, che è un elemento dello spazio l_2 [si è visto a pag. 127 che la a) ivi considerata è vera per ogni $x\in S$]:

$$F : x\in S \longrightarrow (x,e_n)_{n\in N} \in l_2 .$$

Per ogni $x\in S$, i termini della successione $F(x)$ si chiamano le componenti del vettore x rispetto al sistema ortonormale $(e_n)_{n\in N}$.

La F è ovviamente un operatore lineare di S in l_2, ed è invertibile [se $F(x)=0$, la serie di Fourier di x ha tutti i coefficienti nulli, e perciò x (somma della serie) è necessariamente 0]. Si verifica subito che la F è anche suriettiva: infatti qualunque sia $c=(c_n)_{n\in N}$ elemento di l_2, posto $x=\sum_{n=1}^{\infty} c_n e_n$ per la (6.3) di pag. 126 risulta $F(x)=c$. Pertanto F è un isomorfismo algebrico di S su l_2 [1]. Inoltre, per l'uguaglianza di Parseval, osservando che il primo membro è uguale alla norma in l_2 di $F(x)$, si ha:

$$\|F(x)\|_{l_2} = \|x\|_S \qquad \forall x\in S . \tag{15}$$

Dunque l'operatore F, invertibile e suriettivo, non solo "conserva le operazioni di S" (in quanto è lineare), ma "conserva la norma di S", e quindi "conserva le distanze": si dice perciò che F è un isomorfismo isometrico dello spazio di Hilbert S nello spazio di Hilbert l_2, e l'esistenza di un siffatto isomorfismo si esprime dicendo che gli spazi di Hilbert S ed l_2 sono equivalenti, giacché ciascun elemento x di S si può identificare (non solo dal punto di vista algebrico, ma anche dal punto di vista topologico) con la sua immagine $F(x)$ nello spazio l_2.

Abbiamo anticipato alcune locuzioni che saranno introdotte nel capitolo seguente, allo scopo di riassumere le precedenti conclusioni con il teorema:

(6.6). *Ogni spazio di Hilbert separabile di dimensione infinita è equivalente allo spazio di Hilbert l_2.*

---

(1) Cfr. pag. 62. Si noti che F è l'inverso dell'operatore A introdotto a pag. 126.

Concludiamo questo numero avvertendo che si parla anche di serie di Fourier rispetto ad un "sistema ortogonale". Assegnata una successione ortogonale $(h_n)_{n\in N}$ di vettori non nulli di uno spazio di Hilbert S, tutte le locuzioni relative al sistema ortonormale $(e_n)_{n\in N}$ definito dalla relazione:

$$e_n = \frac{h_n}{\|h_n\|} \qquad \forall n \in N$$

si possono riferire anche al "sistema ortogonale" $(h_n)_{n\in N}$. Perciò per ogni $x \in S$ la serie di Fourier di x rispetto al sistema ortonormale $(e_n)_{n\in N}$:

$$(16) \qquad \frac{(x,h_1)}{\|h_1\|^2} h_1 + \frac{(x,h_2)}{\|h_2\|^2} h_2 + \cdots + \frac{(x,h_n)}{\|h_n\|^2} h_n + \cdots$$

si chiama anche <u>la serie di Fourier di x rispetto al sistema ortogonale</u> $(h_n)_{n\in N}$, e i termini della successione:

$$\left(\frac{(x,h_n)}{\|h_n\|^2} \right)_{n\in N}$$

si chiamano <u>i coefficienti di Fourier di x rispetto al sistema ortogonale</u> $(h_n)_{n\in N}$. Il sistema ortogonale $(h_n)_{n\in N}$ si dice poi <u>completo</u> quando non esiste alcun vettore di $S-\{0\}$ che sia ortogonale a ciascun elemento della successione $(h_n)_{n\in N}$.

<u>Esempi</u>.

A puro titolo informativo riportiamo alcuni esempi. Ometteremo le dimostrazioni, tranne quella della completezza del sistema ortogonale dell'es. 1), che sarà esposta a pag. 132.1 -

1) Nello spazio $S = L^2(]-\pi,\pi[)$ la successione:

$$(17) \qquad 1, \cos t, \operatorname{sen} t, \cos 2t, \operatorname{sen} 2t, \ldots, \cos nt, \operatorname{sen} nt, \ldots$$

è ortogonale. Per ogni $x \in S$, i corrispondenti coefficienti di Fourier:

$$\frac{a_0}{2}, a_1, b_1, a_2, b_2, \ldots, a_n, b_n, \ldots$$

sono dati da:

$$a_n = \frac{1}{\pi} \int_{-\pi}^{\pi} x(t) \cos nt \, dt \,, \qquad b_n = \frac{1}{\pi} \int_{-\pi}^{\pi} x(t) \operatorname{sen} nt \, dt \,.$$

Il sistema ortogonale (17) è completo.

È opportuno avvertire che per ogni $x \in S$, il fatto che la serie di Fourier di x rispetto al sistema ortogonale (17) ha per somma x <u>non significa</u> che tale serie, come serie di funzioni, converge per ogni $t \in]-\pi,\pi[$ ed ha per somma $x(t)$; significa invece, ricordando la norma in $L^2(]-\pi,\pi[)$ e denotando con $s_n(t)$ la somma parziale n^{ma} della serie, che risulta:

$$\lim_{n\to\infty} \int_{-\pi}^{\pi} [x(t) - s_n(t)]^2 \, dt = 0 \,,$$

ossia che $s_n(t)$ converge ad $x(t)$ in media quadratica. Com'è noto, esistono vari

criteri di convergenza puntuale e di convergenza uniforme, che sono oggetto di studio nella teoria delle "serie trigonometriche di Fourier"; con tale locuzione si intendono appunto le serie di Fourier degli elementi di $L^2(]-\pi,\pi[)$ rispetto al sistema ortogonale (17), che peraltro vengono spesso chiamate semplicemente "serie di Fourier".

Naturalmente l'intervallo $]-\pi,\pi[$ può essere sostituito con un qualunque intervallo di ampiezza 2π.

2) Nello spazio $L^2(]-1,1[)$, ortonormalizzando la successione di vettori $(x_n)_{n\in N}$ definita da:

$$x_n(t) = t^{n-1} \qquad \forall t \in]-1,1[\tag{18}$$

si ottiene un sistema ortonormale completo $(e_n)_{n\in N}$. I termini di tale successione, che ovviamente sono polinomi, si chiamano <u>polinomi di Legendre</u>, e si incontrano nella valutazione asintotica di particolari soluzioni di certe equazioni differenziali dipendenti da un parametro.

3) Ortonormalizzando la successione (18) nello spazio $L^2(]-1,1[, \mu)$ con

$$d\mu = \frac{1}{\sqrt{1-t^2}}dt \qquad \left[\, d\mu = \sqrt{1-t^2}\,dt \,\right] \qquad (1)$$

si ottiene un sistema ortonormale completo $(e_n)_{n\in N}$. I termini di tale successione si chiamano i <u>polinomi di Tchebicev di prima specie</u> [<u>di seconda specie</u>]

4) Ortonormalizzando la successione (18) nello spazio $L^2(R, \mu)$ con $d\mu = e^{-t^2}dt$, si ottengono i cosiddetti <u>polinomi di Hermite</u>, che costituiscono un sistema ortonormale completo.

5) Ortonormalizzando la successione (18) nello spazio $L^2(R^+, \mu)$ con $d\mu = e^{-t}dt$, si ottiene un sistema ortonormale completo (<u>polinomi di Laguerre</u>).

---

(1) Cfr. es. 3 pag. 55 e es. 3 pag. 100. E' opportuno ricordare quanto detto nel n. 7 di pag. 45.

Ritornando sull'esempio 1), dimostriamo che:

(6.7). Il sistema ortogonale (17) [pag. 131] è completo.

Si tratta di stabilire che, se un vettore x di $L^2(]-\pi,\pi[)$ è ortogonale al sottospazio generato dal sistema (17), risulta necessariamente $x=0$.

In primo luogo poniamo:

$$x_1(t)=\int_0^t x(\tau)d\tau \qquad \forall t\in[-\pi,\pi],$$

e teniamo presente che, per l'ipotesi su x, i coefficienti di Fourier di x (rispetto al sistema ortogonale (17)) sono tutti nulli; in particolare dal fatto che è nullo il primo si deduce l'uguaglianza $x_1(-\pi)=x_1(\pi)$.

Se allora si considerano i coefficienti di Fourier di x_1, escluso il primo, e si calcolano mediante integrazione per parti col fattore finito x_1 (lecita in quanto x_1 è assolutamente continua), tenuto conto che $x_1'=x$ q.o. si trova che essi sono tutti nulli.

Tale conclusione comporta che il ragionamento fatto per x_1 si può ripetere per la funzione:

$$x_2(t)=\int_0^t x_1(\tau)d\tau+ct \quad \forall t\in[-\pi,\pi], \text{ con } c=\frac{-1}{2\pi}\int_{-\pi}^{\pi}x_1(\tau)d\tau$$

la quale, per la scelta di c, verifica l'uguaglianza:

$$x_2(-\pi)=x_2(\pi); \tag{19}$$

occorre ora tener conto che $x_2'=x_1+c$, e osservare che la costante c, come elemento di $L^2(]-\pi,\pi[)$, è ortogonale a tutti i vettori (17) escluso il primo.

Come già per x_1, si deduce allora che i coefficienti di Fourier di x_2, a partire dal secondo, sono tutti nulli: dunque la somma della serie di Fourier di x_2 è costante. D'altra parte, com'è noto, siccome $x_2\in C^1([-\pi,\pi])$ (in quanto $x_2'=x_1+c$ e x_1 è continua) e sussiste la (19), x_2 coincide con la somma della sua serie di Fourier: dunque x_2 è costante.

Pertanto $x_2''=0$: siccome $x_2'=x_1+c$ e quindi $x_2''=x_1'=x$ q.o., si conclude che x è il vettore nullo di $L^2(]-\pi,\pi[)$ e l'assunto è provato.

APPENDICE AL n. 5 di pag. 98

Dimostrazione di alcuni risultati sullo spazio l_p.

(5.1). Lo spazio l_p è completo.

Per ogni $n \in N$ sia

$$x^{(n)} = (x_1^{(n)}, x_2^{(n)}, \dots, x_k^{(n)}, \dots)$$

un punto di l_p, e la successione:

(1) $$x^{(1)}, x^{(2)}, \dots, x^{(n)}, \dots$$

sia una successione di Cauchy di l_p. Si tratta di dimostrare che la (1) è convergente in l_p, e a tale scopo fissiamo un numero positivo ε.

Per l'ipotesi, esiste $\nu \in N$ tale che si abbia:

$$\| x^{(r+m)} - x^{(r)} \| < \frac{\varepsilon}{4} \quad \forall r > \nu \text{ e } \forall m \in N,$$

ossia:

(2) $$\sum_{i=1}^{\infty} | x_i^{(r+m)} - x_i^{(r)} |^p < \left(\frac{\varepsilon}{4}\right)^p \quad \forall r > \nu \text{ e } \forall m \in N.$$

Poiché ciascun termine della serie a primo membro è maggiorato dal secondo membro, ne segue che qualunque sia $i \in N$ la successione:

(3) $$x_i^{(1)}, x_i^{(2)}, \dots, x_i^{(n)}, \dots$$

è una successione di Cauchy di numeri reali o complessi, quindi è convergente in C. Detto λ_i il limite della (3), proviamo che la successione:

(4) $$\lambda = (\lambda_1, \lambda_2, \dots, \lambda_n, \dots)$$

appartiene allo spazio l_p ed è il limite in l_p della (1).

Infatti dalla (2) si deduce, per ogni $k \in N$:

$$\sum_{i=1}^{k} | x_i^{(r+m)} - x_i^{(r)} |^p < \left(\frac{\varepsilon}{4}\right)^p \quad \forall r > \nu \text{ e } \forall m \in N$$

e quindi, passando al limite per $m \to \infty$:

(5) $$\sum_{i=1}^{k} | \lambda_i - x_i^{(r)} |^p \leq \left(\frac{\varepsilon}{4}\right)^p < \left(\frac{\varepsilon}{2}\right)^p \quad \forall r > \nu \ ;$$

essendo $|\lambda_i|^p \leq |\lambda_i - x_i^{(r)} + x_i^{(r)}|^p \leq 2^p \left(|\lambda_i - x_i^{(r)}|^p + |x_i^{(r)}|^p \right)$, ne segue:

$$\sum_{i=1}^{k} |\lambda_i|^p \leq \varepsilon^p + \| x^{(r)} \| \qquad \forall r > \nu.$$

Scrivendo tale relazione per un fissato $n > \nu$, ad es. per $n = \nu+1$, se ne deduce che la successione (4) è un elemento dello spazio l_p. Dopo ciò, poiché la (5) implica:

$$\sum_{i=1}^{\infty} |\lambda_i - x_i^{(n)}|^p \leq \left(\frac{\varepsilon}{2}\right)^p < \varepsilon^p \qquad \forall n > \nu,$$

ossia:

$$\|\lambda - x^{(n)}\| < \varepsilon \qquad \forall n > \nu,$$

si conclude che $\lambda = \lim x^{(n)}$ e l'assunto è provato.

(5.2). La famiglia di spazi $(l_p)_{p \in [1,+\infty[}$ è crescente in senso stretto rispetto alla relazione di inclusione. Si ha cioè:

$$p < q \implies l_p \subset l_q .$$

Qualunque sia $x \in l_p$, $x = (x_1, \dots, x_n, \dots)$, dalla convergenza della serie $\sum |x_i|^p$ consegue che la successione $(x_i)_{i \in N}$ è infinitesima, e quindi esiste $\nu \in N$ tale da risultare $|x_i| < 1$ $\forall i > \nu$. Per ogni $i > \nu$ si ha allora, supposto $p < q$:

$$|x_i|^q < |x_i|^p$$

e perciò, essendo convergente la serie $\sum |x_i|^p$, anche la serie $\sum |x_i|^q$ è convergente. Dunque $x \in l_q$ e ciò prova che $l_p \subseteq l_q$.

Per dimostrare l'inclusione stretta basta scegliere un numero α tale che

$$\frac{1}{q} < \alpha < \frac{1}{p}$$

e osservare che la successione $\left(\frac{1}{n^\alpha}\right)_{n \in N}$ appartiene ad l_q ma non appartiene ad l_p.

(5.3) Siano:

$$x^{(n)} = (x_1^{(n)}, x_2^{(n)}, \dots, x_k^{(n)}, \dots) \qquad n \in N$$

$$x^{(0)} = (x_1^{(0)}, x_2^{(0)}, \dots, x_k^{(0)}, \dots)$$

punti di l_2, e per ogni $k \in N$ si ponga:

$$P_k^{(n)} = (x_1^{(n)}, x_2^{(n)}, \dots, x_k^{(n)})$$

$$P_k^{(0)} = (x_1^{(0)}, x_2^{(0)}, \dots, x_k^{(0)}).$$

Sussiste l'equivalenza:

$\lim x^{(n)} = x^{(0)}$ in $l_2 \iff \lim_n P_k^{(n)} = P_k^{(0)}$ in R^k, uniformemente rispetto a k

Se $x^{(n)} \to x^{(0)}$ in l_2, per ogni $\varepsilon > 0$ esiste $\nu \in N$ tale da risultare:

$$\sum_{i=1}^{\infty} |x_i^{(n)} - x_i^{(0)}|^2 < \varepsilon^2 \qquad \forall n > \nu.$$

Di conseguenza qualunque sia $k \in N$ risulta:

$$\sum_{i=1}^{k} |x_i^{(n)} - x_i^{(0)}|^2 < \varepsilon^2 \qquad \forall n > \nu,$$

ossia:

$$\| P_k^{(n)} - P_k^{(0)} \|_{R^k} < \varepsilon \qquad \forall n > \nu.$$

Viceversa, supponiamo che $P_k^{(n)}$ converga in R^k a $P_k^{(0)}$ uniformemente rispetto a k. Per ogni $\varepsilon > 0$ esiste allora $\nu \in N$ tale che si abbia:

$$\| P_k^{(n)} - P_k^{(0)} \|_{R^k} < \frac{\varepsilon}{2} \qquad \forall n > \nu \text{ e } \forall k \in N,$$

ossia:

$$\sum_{i=1}^{k} |x_i^{(n)} - x_i^{(0)}|^2 < \frac{\varepsilon^2}{4} \qquad \forall n > \nu \text{ e } \forall k \in N.$$

Passando al limite per $k \to \infty$ ne segue che:

$$\sum_{i=1}^{\infty} |x_i^{(n)} - x_i^{(0)}|^2 \le \frac{\varepsilon^2}{4} < \varepsilon^2 \qquad \forall n > \nu$$

ossia:

$$\| x^{(n)} - x^{(0)} \|_{l_2} < \varepsilon \qquad \forall n > \nu.$$

Pertanto $x^{(n)} \to x^{(0)}$ in l_2 e l'asserto è provato.

Capitolo VI – Operatori lineari e continui tra spazi vettoriali normati . Funzionali lineari e continui.

1. Applicazioni continue . Continuità uniforme. Isometrie.

Siano : S, S_1 spazi topologici , f un'applicazione di $X \subseteq S$ in S_1 , x_0 un punto di X.

Come abbiamo già visto nel Cap. I, si dice che f é continua nel punto x_0 quando é vera la proposizione seguente:

(*) Per ciascun intorno I' di $f(x_0)$ (in S_1), esiste un intorno I di x_0 in X tale da verificare la seguente implicazione ,qualunque sia $x \in X$:

$$x \in I \implies f(x) \in I'.$$

La proposizione (*) si può esprimere equivalentemente in forma diversa se uno degli spazi S, S_1 é uno spazio metrico, o vettoriale topologico, o vettoriale normato, o pre-hilbertiano. Pur essendo ciò ovvia conseguenza di quanto detto nel Cap. IV, vogliamo esporre, per comodità del lettore, le diverse formulazioni della (*) nei vari casi.

Con riferimento allo spazio S , nella (*) si possono sostituire le parole " esiste un intorno I di x_0 " , e successivamente la proprietà " $x \in I$ ", con quelle indicate nel seguente schema:

	$\exists$ un intorno I di x_0	$x \in I$
$S=(\Sigma,d)$ sp. metrico	$\exists\, \delta>0$	$d(x,x_0)<\delta$
S sp. vett. topologico	$\exists$ un intorno J di 0	$x \in x_0+J$
S sp. vett. normato	$\exists\, \delta>0$	$\|x-x_0\|<\delta$
S sp. pre-hilbertiano	$\exists\, \delta>0$	$(x-x_0,x-x_0)<\delta$

Analogamente, con riferimento allo spazio S_1 , nella (*) si possono sostituire le parole " per ciascun intorno I' di $f(x_0)$ " , e successivamente la proprietà " $f(x) \in I'$ " , con quelle indicate nel seguente schema:

	$\forall$ intorno $\mathcal{I}'$ di $f(x_0)$	$f(x)\in\mathcal{I}'$
$S_1=(\Sigma_1,d_1)$ sp. metrico	$\forall\varepsilon>0$	$d_1(f(x),f(x_0))<\varepsilon$
S_1 sp.vett.topologico	$\forall$ intorno $\mathcal{J}'$ di 0	$f(x)\in f(x_0)+\mathcal{J}'$
S_1 sp.vett.normato	$\forall\varepsilon>0$	$\|f(x)-f(x_0)\|<\varepsilon$
S_1 sp.pre-hilbertiano	$\forall\varepsilon>0$	$(f(x)-f(x_0),f(x)-f(x_0))<\varepsilon$

Ad esempio, se S e S_1 sono entrambi spazi vettoriali normati, la $(*)$ si può formulare equivalentemente come segue:

$(**)$ $\quad \forall\varepsilon>0\ \exists\delta>0$ tale che, qualunque sia $x\in X$:
$\|x-x_0\|_S<\delta \implies \|f(x)-f(x_0)\|_{S_1}<\varepsilon$.

Negli schemi sopra riportati abbiamo esaminato anche il caso in cui S o S_1 sia uno spazio pre-hilbertiano, solo per completare gli schemi stessi: in tal caso, ovviamente, è più semplice considerare S o S_1 come spazi vettoriali normati.

E' opportuno quì ricordare che, nel caso in cui S é uno spazio metrico (in particolare, spazio vettoriale normato o spazio pre-hilbertiano), la continuità di f in x_0 si può esprimere equivalentemente come "continuità per successioni", in virtù del teorema (1.10) di pag.110.

Ricordiamo inoltre che, qualunque siano gli spazi topologici S, S_1, la continuità di f in X si può esprimere utilizzando soltanto la nozione di insieme aperto o quella di insieme chiuso [cfr. (7.4) pag.22].

Pur essendo ovvio, é utile osservare che:

(1) (S,d) spazio metrico $\implies$ d funzione continua in $S\times S$

(2) (S,ν) spazio vett.normato $\implies$ ν funzione continua in S

(3) (S,p) spazio pre-hilbertiano $\implies$ p funzione continua in $S\times S$.

DIM.(1). Ricordiamo che la distanza d^* di due punti di $S\times S$ é definita ponendo, qualunque siano gli elementi $z=(x,y)$ e $z_0=(x_0,y_0)$ di $S\times S$:

$$d^*(z,z_0)=d(x,x_0)+d(y,y_0)\,.$$

Ciò premesso, dalle relazioni:

$$d(x,y)\le d(x,x_0)+d(x_0,y) \quad , \quad d(x_0,y)\le d(x_0,y_0)+d(y_0,y)$$

Consegue:

$$d(x,y)-d(x_0,y_0) \leq d(x,x_0)+d(y,y_0) = d^*(z,z_0)$$

Scambiando x con x_0 e y con y_0, si ha che l'opposto del primo membro della soprascritta relazione é $\leq$ di $d^*(z,z_0)$; pertanto:

$$|d(x,y)-d(x_0,y_0)| \leq d^*(z,z_0) .$$

Dopo ciò, l'asserto é banale .

Dim. (2) . Basta osservare che, qualunque siano $x, x_0 \in S$, si ha :

$$\|x\| \leq \|x-x_0\|+\|x_0\| \quad , \qquad \|x_0\| \leq \|x_0-x\|+\|x\|$$

e quindi:

$$\big|\, \|x\|-\|x_0\| \,\big| \leq \|x-x_0\| .$$

Dim. (3). Sia $(x_0,y_0) \in S \times S$. Per ogni $\varepsilon > 0$, sia δ un numero positivo , minore di 1 e tale che:

$$\delta < \varepsilon / (\|x_0\|+\|y_0\|+1) .$$

Qualunque sia $(x,y) \in S \times S$ tale che $\|x-x_0\| < \delta$ e $\|y-y_0\| < \delta$ si ha allora, tenendo presente la disuguaglianza di Schwarz [cfr.(5) pag.97] :

$$|p(x,y)-p(x_0,y_0)| \leq |p(x-x_0,y)|+|p(x_0,y-y_0)| \leq$$
$$\leq \|x-x_0\| \cdot \|y\| + \|x_0\| \cdot \|y-y_0\| < \delta\,(\|y_0\|+\delta) + \|x_0\|\,\delta <$$
$$< \delta\,(\|y_0\|+1) + \|x_0\|\,\delta < \varepsilon .$$

Introduciamo ora la nozione di continuità uniforme, nel caso in cui S e S_1 sono spazi metrici (in particolare vettoriali normati o pre-hilbertiani).

Siano (S,d), (S_1,d_1) spazi metrici, X un sottoinsieme di S, ed f un'applicazione di X in S_1. Si dice che f é uniformemente continua (in X), quando per ciascun $\varepsilon > 0$ esiste $\delta > 0$ tale da verificare la seguente implicazione, qualunque siano $x_1, x_2 \in X$:

$$d(x_1,x_2) < \delta \Rightarrow d_1(f(x_1),f(x_2)) < \varepsilon .$$

Sussiste il teorema (teorema di Cantor):

(1.1) Siano S, S_1 spazi metrici, X un sottoinsieme di S, ed f un'applicazione di S in S_1. Si ha che:

(f continua e X compatto) $\Rightarrow$ f uniformemente continua.

Dim. Procedendo per assurdo, ammettiamo che, essendo f continua e X compatto, la f non sia uniformemente continua. Esiste allora $\varepsilon > 0$ tale che, per ogni $n \in N$, esistono $x_n, y_n \in X$ soddisfacenti alle relazioni:

(4) $\quad d(x_n, y_n) < \frac{1}{n} \quad , \quad d_1(f(x_n), f(y_n)) \geq \varepsilon$ (1).

Essendo X compatto, ossia [cfr. (2.1) pag.111] sequenzialmente compatto, dalla successione $(x_n)_{n \in N}$ si può estrarre una successione, diciamola $(x_{k_n})_{n \in N}$, convergente ad un punto di X, diciamolo x_0. Risultando, per la prima delle (4):

$$d(x_{k_n}, y_{k_n}) < \frac{1}{k_n} \qquad \forall n \in N$$

anche la successione $(y_{k_n})_{n \in N}$ converge al punto x_0. In virtù della (1.10) di pag.110 si ha allora:

$$\lim_{n \to \infty} f(x_{k_n}) = \lim_{n \to \infty} f(y_{k_n}) = f(x_0)$$

e quindi

$$\lim_{n \to \infty} d_1(f(x_{k_n}), f(y_{k_n})) = 0 \ .$$

Ma ciò è assurdo, in quanto contraddice la seconda delle (4).

La (1.1.) esprime, nel caso S, S_1 spazi metrici, una proprietà delle funzioni continue in un insieme compatto; è opportuno ricordare le proprietà di dette funzioni (7.5), (7.6) pag.22-23, che sussistono con S, S_1 spazi topologici qualunque, nonchè la (7.7) di pag.23 relativa al caso $S_1 = R$.

---

(1) Come al solito, abbiamo indicato con d la metrica in S, e con d_1 la metrica in S_1.

Concludiamo questo numero con la seguente definizione.

Siano (S,d) , (S_1,d_1) spazi metrici, ed f un'applicazione di $X\subseteq S$ in S_1. Si dice che f é una <u>isometria</u> quando risulta :

$$(5)\qquad d_1(f(x_1),f(x_2)) = d(x_1,x_2) \qquad \forall x_1,x_2 \in X.$$

Nel caso in cui S e S_1 sono spazi vettoriali normati, la (5) si scrive:

$$(5')\qquad \|f(x_1)-f(x_2)\|_{S_1} = \|x_1-x_2\|_S \qquad \forall x_1,x_2\in X\ ;$$

osserviamo che, se A é un operatore <u>lineare</u> di S in S_1, il fatto che A é un'isometria [cioé che vale la (5') con $f=A$] equivale alla relazione:

$$(6)\qquad \|Ax\|_{S_1} = \|x\|_S \qquad \forall x\in S.$$

Pur essendo banale, rileviamo che le isometrie sono funzioni uniformemente continue.

2. <u>Operatori lineari e continui tra spazi vettoriali normati. Isomorfismi.</u>

Siano S, S_1 spazi vettoriali normati, ed A un operatore lineare di S in S_1.

Si dice che A e' un operatore <u>limitato</u>, quando esiste un numero positivo l tale che:

$$(1)\qquad \|Ax\| \le l\,\|x\| \qquad \forall x\in S$$

La parola "limitato" e' dovuta al fatto che la (1) equivale alla proprietà dell'operatore A di trasformare insiemi limitati di S in insiemi limitati di S_1; precisamente, denotando con X un sottoinsieme di S, la (1) equivale all'implicazione:

$$(2)\qquad \underline{X \text{ limitato}} \Rightarrow \underline{A(X) \text{ limitato}}.$$

Dim. (1) $\Rightarrow$ (2). Se X e' limitato, esiste un numero positivo k tale che $\|x\|\le k$ $\forall x\in X$; e quindi per ogni $x\in X$ dalla (1) consegue $\|Ax\|\le lk$: pertanto $A(X)$ e' limitato.

Dim. (2) $\Rightarrow$ (1). Se vale la (2), l'immagine mediante A della sfera chiusa di S di centro 0 e raggio 1 e' un insieme limitato; dunque esiste $l>0$ tale che per $y\in S$ si abbia:

$$(3)\qquad \|y\|\le 1 \Rightarrow \|Ay\|\le l.$$

Qualunque sia $x\in S-\{0\}$, dalla (3) per $y=\frac{x}{\|x\|}$ consegue $\|A\left(\frac{x}{\|x\|}\right)\|\le l$, ossia $\frac{1}{\|x\|}\|Ax\|\le l$, donde la (1) per $x\neq 0$. Poiché la (1) sussiste banalmente anche per $x=0$, l'asserto e' provato.

Se un operatore lineare e' limitato, l'operatore e' anche continuo, e viceversa; per gli operatori lineari la continuità (1) si può equivalentemente esprimere anche in diversi altri modi, come risulta dal seguente teorema:

---

(1) La locuzione "operatore lineare" si adoperava talvolta per intendere "operatore lineare e continuo": peraltro l'abitudine poco opportuna e' caduta in disuso.

(2.1) Se A e' un operatore lineare di S in S_1, le proposizioni che seguono sono equivalenti:

c_1) A e' continuo in S.

c_2) Esiste almeno un punto di S in cui A e' continuo.

c_3) Esiste $l>0$ tale che: $\|x\| \leq 1 \Rightarrow \|Ax\| \leq l$.

c_4) Esiste $l>0$ tale che: $\|x\| = 1 \Rightarrow \|Ax\| \leq l$.

c_5) Esiste $l>0$ tale che: $\|Ax\| \leq l\|x\| \quad \forall x \in S$, ossia A e' limitato.

c_6) A e' una funzione lipschitziana, ossia esiste $l>0$ tale che $\|Ax-Ay\| \leq l\|x-y\| \quad \forall x,y \in S$.

c_7) A e' una funzione uniformemente continua.

Dim.

$c_1 \Rightarrow c_2$. Evidente.

$c_2 \Rightarrow c_3$. Se A e' continuo nel punto x_0, esiste $\delta>0$ tale che per $x \in S$ si abbia:

$$\|x-x_0\| \leq \delta \Rightarrow \|Ax-Ax_0\| < 1 ,$$

cioè:

$$\|x-x_0\| \leq \delta \Rightarrow \|A(x-x_0)\| < 1 .$$

Ovviamente qualunque punto y di S si può esprimere nella forma $x-x_0$ (basta porre $x = x_0+y$), e quindi la precedente implicazione si può scrivere:

$$(*) \qquad \|y\| \leq \delta \Rightarrow \|Ay\| < 1 .$$

Ciò premesso, sia $x \in S$ tale che $\|x\| \leq 1$. Essendo $\|x\delta\| = \delta\|x\| \leq \delta$, dalla $(*)$ segue $\|A(x\delta)\| < 1$, ossia $\delta\|Ax\| < 1$: pertanto e' vera la c_3 con $l = 1/\delta$.

$c_3 \Rightarrow c_4$. Evidente.

$c_4 \Rightarrow c_5$. Qualunque sia $x \in S-\{0\}$, poichè il vettore $\dfrac{x}{\|x\|}$ ha norma uguale ad 1, per la c_4) si ha:

$$\left\| A\left(\frac{x}{\|x\|}\right) \right\| \leq l .$$

e quindi $\|Ax\| \leq l\|x\|$. Poichè tale relazione e' vera anche per $x=0$, la c_5) e' provata.

$c_5 \Rightarrow c_6$. Basta utilizzare la c_5) con $x-y$ in luogo di x.

$c_6 \Rightarrow c_7$. Evidente.

$c_7 \Rightarrow c_1$. Evidente.

Abbiamo visto che la continuità di un operatore lineare $A: S \to S$, equivale all'esistenza di un numero positivo l tale da risultare:

$$(4) \qquad \|Ax\| \leq l\|x\| \qquad \forall x \in S .$$

Orbene, se esiste un numero $l>0$ per il quale in luogo della (4) si verifica la disuguaglianza opposta, l'operatore A e' invertibile e il suo inverso A^{-1} e' continuo. Infatti

sussiste il teorema:

(2.2) Se A è un operatore lineare di S in S_1, le seguenti proposizioni sono equivalenti:

i_1) Esiste $l'>0$ tale che $\|Ax\| \geq l'\|x\| \quad \forall x \in S$.

i_2) L'operatore A è invertibile, e l'inverso A^{-1} è continuo.

Dim.

$i_1 \Rightarrow i_2$. Dalla disuguaglianza:

$$(*) \qquad \|x\| \leq \frac{1}{l'}\|Ax\| \qquad \forall x \in S$$

consegue che sussiste l'implicazione:

$$Ax = 0 \Rightarrow x = 0,$$

ossia che A è invertibile (pag.72-73). D'altra parte per ogni $y \in A(S)$, essendo $A^{-1}y$ un punto di S risulta, per la $(*)$:

$$\|A^{-1}y\| \leq \frac{1}{l'}\|AA^{-1}y\| = \frac{1}{l'}\|y\| .$$

Pertanto l'operatore A^{-1} è limitato, ossia continuo.

$i_2 \Rightarrow i_1$). Essendo A^{-1} continuo, esiste $l_1 > 0$ tale che:

$$\|A^{-1}y\| \leq l_1\|y\| \qquad \forall y \in A(S),$$

e quindi risulta:

$$\|A^{-1}Ax\| \leq l_1\|Ax\| \qquad \forall x \in S,$$

ossia:

$$\|x\| \leq l_1\|Ax\| \qquad \forall x \in S.$$

Pertanto è vera la i_1) con $l' = 1/l_1$.

Se A è un operatore lineare di S in S_1, per le (2.1), (2.2) l'esistenza di due numeri positivi l, l' tali da risultare:

$$(5) \qquad l'\|x\| \leq \|Ax\| \leq l\|x\| \qquad \forall x \in S,$$

equivale al fatto che A è invertibile ed è continuo insieme al suo inverso A^{-1} [che è definito nel sottospazio $A(S)$ di S_1]. Si dice perciò che A è un isomorfismo dello spazio vettoriale normato S sul sottospazio vettoriale normato $A(S)$, e in particolare, se $A(S) = S_1$ (cioè se A è suriettivo), si dice che A è un isomorfismo dello spazio vettoriale normato S sullo spazio vettoriale normato S_1.

Dunque un isomorfismo di S su S_1 è un isomorfismo algebrico per il quale esistono due numeri positivi l, l' tali da verificare la (5), cioè un isomorfismo algebrico che sia continuo insieme al suo inverso (*). Notiamo che la (5) è in particolare verificata se l'operatore lineare A è una isometria [cfr. (6) pag.137]; un isomorfismo algebrico A

(*) La nozione di isomorfismo algebrico è stata introdotta a pag.62.

che sia anche una isometria si chiama perciò un isomorfismo isometrico dello spazio vettoriale normato S sul sottospazio vettoriale normato $A(S)$, e in particolare su S_1 se $A(S) = S_1$. Un esempio di isomorfismo isometrico è l'operatore F considerato a pag. 130.

Non è detto che, qualunque siano gli spazi vettoriali normati S, S_1, esista un isomorfismo di S su S_1. Quando esiste un isomorfismo di S su S_1 si dice che gli spazi vettoriali normati S ed S_1 sono isomorfi.

In particolare, se esiste un isomorfismo isometrico A di S su S_1, l'operatore A è una corrispondenza biunivoca di S su S_1 che non solo conserva le operazioni di addizione e di moltiplicazione per uno scalare (1), ma "conserva anche la norma" in quanto risulta:

$$\|Ax\|_{S_1} = \|x\|_S \qquad \forall x \in S .$$

Si dice perciò che gli spazi S ed S_1 sono equivalenti, e in tal caso ciascun elemento x di S si può identificare con il suo corrispondente Ax in S_1.

Se esiste un isomorfismo A di S su S_1 non isometrico, e si identifica ogni elemento x di S con il suo corrispondente Ax in S_1, la norma in S_1 può essere considerata come un'altra norma per il sostegno di S, equivalente alla norma in S in virtù della (5). Con linguaggio impreciso ma espressivo, si può dire allora che, se S ed S_1 sono spazi vettoriali normati isomorfi, si può "trasportare" nello spazio S_1 tutto ciò che si verifica nello spazio S, con riferimento non solo a proprietà di carattere algebrico e topologico, ma anche a proprietà derivanti dalla norma; se poi S ed S_1 sono equivalenti, gli spazi S ed S_1 si possono considerare addirittura come il medesimo spazio.

Osserviamo ora che:

(2.3) Se l'operatore lineare A di S in S_1 è un isomorfismo di S su $A(S)$, e se S è uno spazio di Banach, il sottospazio $A(S)$ di S_1 è chiuso in S_1.

Dim. Ricordando la (1.8) di pag. 109 si tratta di far vedere che, se $(y_n)_{n\in N}$ è una successione convergente di punti di $A(S)$, il suo limite y_0 appartiene ad $A(S)$.

Posto $x_n = A^{-1}y_n$, qualunque siano $p, q \in N$ si ha per la (5):

$$(6) \qquad \ell' \|x_p - x_q\| \leq \|A(x_p - x_q)\| = \|y_p - y_q\| .$$

Essendo $(y_n)_{n\in N}$ una successione di Cauchy in quanto convergente, per la (6) anche $(x_n)_{n\in N}$ è una successione di Cauchy: pertanto, siccome S è uno spazio di Banach, la $(x_n)_{n\in N}$ è convergente. Posto $\lim x_n = x_0$, per la continuità dell'operatore A risulta:

$$y_0 = \lim y_n = \lim Ax_n = Ax_0$$

e quindi $y_0 \in A(S)$.

---

(1) Ricordiamo che con tale locuzione, che a questo punto appare significativa, abbiamo espresso il fatto che A è lineare.

Supposto che A sia un isomorfismo algebrico del sostegno di S sul sostegno di S_1 [brevemente, diremo un isomorfismo algebrico di S su S_1], se S ed S_1 sono spazi di Banach per verificare che A è un isomorfismo dello spazio di Banach S sullo spazio di Banach S_1 è sufficiente stabilire che A è continuo, perchè la continuità dell'operatore inverso A^{-1} discende in conseguenza. Sussiste infatti il seguente notevole teorema di Banach, del quale omettiamo la dimostrazione:

(2.4) *Siano S, S_1 spazi di Banach, ed A un isomorfismo algebrico di S su S_1. Se l'operatore A è continuo, anche l'inverso A^{-1} è continuo.*

Naturalmente il teorema di Banach si può applicare anche ad un isomorfismo algebrico di S su $A(S)$ [ossia, ad un operatore lineare invertibile, non necessariamente suriettivo], ma in tal caso occorre aggiungere l'ipotesi che $A(S)$ sia chiuso in S_1 [1], per assicurare che il sottospazio di S_1 di sostegno $A(S)$ sia uno spazio di Banach [teor. (3.2) pag. 112].

E' importante notare che:

(2.5) *Ogni spazio vettoriale normato E, di dimensione finita n, è isomorfo allo spazio di Banach K^n* [2].

Fissata una base $\{e_1, \dots, e_n\}$ di E, proveremo che l'operatore A definito come segue:

$$A: (\lambda_1, \dots, \lambda_n) \in K^n \to \sum_{i=1}^{n} \lambda_i e_i \in E \tag{7}$$

è un isomorfismo dello spazio normato K^n sullo spazio normato E (*isomorfismo canonico di K^n su E*).

Dim. Ovviamente A è lineare, ed è invertibile perchè $A\lambda = 0 \Rightarrow \lambda = 0$ [cfr. (2') pag. 73].

Poichè A è anche suriettivo (per definizione di base), ne segue che A è un isomorfismo algebrico di K^n su E.

D'altra parte, detto h il più grande dei numeri $\|e_1\|, \dots, \|e_n\|$, per ogni $\lambda = (\lambda_1, \dots, \lambda_n) \in K^n$ si ha:

$$\|A\lambda\| = \left\| \sum_{i=1}^{n} \lambda_i e_i \right\| \leq \sum_{i=1}^{n} |\lambda_i| \cdot \|e_i\| \leq h \sum_{i=1}^{n} |\lambda_i| = h \sum_{i=1}^{n} \sqrt{\lambda_i^2} \leq h \cdot n \|\lambda\|$$

sicchè A è limitato, ossia continuo.

Se si suppone che E sia uno spazio di Banach, dal teorema di Banach (2.4) consegue che A è un isomorfismo dello spazio E nello spazio di Banach R^n, donde l'assurdo. Ma nel teorema in esame lo spazio E non è supposto completo, e quindi la continuità dell'operatore

---

[1] E' opportuno notare che non è applicabile il precedente teorema (2.3), giacchè ivi l'operatore A è un isomorfismo di S su $A(S)$ non soltanto algebrico. D'altronde, se si supponesse A isomorfismo tra spazi normati, la tesi del teorema di Banach sarebbe già compresa nell'ipotesi.

[2] Quindi ogni spazio vettoriale normato a dimensione finita è uno spazio di Banach. In proposito è utile quanto detto nella nota [1] di pag. 203 BIS.

inverso A^{-1} va dimostrata direttamente.

Detto Γ l'insieme dei punti $\lambda \in K^n$ tali che $\|\lambda\| = 1$, consideriamo la funzione reale φ di n variabili reali definita come segue:

$$\varphi : \lambda = (\lambda_1, \dots, \lambda_n) \in \Gamma \to \|A\lambda\| \in R_0^+ .$$

Essendo φ continua [poiché composta mediante le funzioni continue A e $\forall x \in E \to \|x\|_E$ (cfr. (2) pag. 134)] e Γ compatto, la φ è dotata di minimo, diciamolo m. Si ha allora:

$$\|A\lambda\| \geq m \qquad \forall \lambda \in \Gamma , \tag{8}$$

e il numero m è certamente positivo: infatti non può essere $m = 0$, altrimenti esisterebbe $\lambda = (\lambda_1, \dots, \lambda_n) \in \Gamma$ tale che $A\lambda = 0$, il che è assurdo in quanto $A\lambda$ è la combinazione lineare dei vettori $e_1, \dots, e_n$ (linearmente indipendenti) di coefficienti $\lambda_1, \dots, \lambda_n$ non tutti nulli (giacché $\|\lambda\| = 1$).

Qualunque sia $\lambda \in K^n - \{0\}$, essendo $\lambda / \|\lambda\|$ un punto di Γ, per la (8) risulta:

$$\left\| A\left(\frac{\lambda}{\|\lambda\|} \right) \right\| \geq m ,$$

ossia:

$$\|A\lambda\| \geq m \|\lambda\|$$

e questa relazione vale banalmente anche per $\lambda = 0$. Per il teor. (2.2) di pag. 139 ne segue che A^{-1} è continua, e ciò prova l'asserto.

— — — — — — — — — — — — —

Se A è un'applicazione di S in S_1, il sottoinsieme di $S \times S_1$:

$$G = \{ (x, y) \in S \times S_1 : y = A(x) \}$$

si chiama <u>il grafico di A</u>. Tale locuzione interviene nel teorema seguente, che è opportuno accostare al teor. (2.1) di pag. 138 in quanto fornisce, nel caso che S ed S_1 siano spazi di Banach, delle ulteriori caratterizzazioni della continuità di un operatore lineare.

(2.6) <u>Se S, S_1 sono spazi di Banach ed A è un operatore lineare di S in S_1, le proposizioni che seguono sono equivalenti:</u>

a) <u>Per ogni successione $(x_n)_{n \in N}$ di punti di S avente limite 0 e tale che la successione $(Ax_n)_{n \in N}$ sia convergente, risulta:</u>

$$\lim Ax_n = 0 . \tag{9}$$

b) <u>Il grafico di A è un insieme chiuso.</u>

c) <u>A è continuo.</u>

La b) si esprime dicendo che A è un <u>operatore chiuso</u>, e l'implicazione $b \Rightarrow c$ viene chiamata <u>teorema del grafico chiuso.</u>

Prima di dimostrare il teorema ci sembra opportuno fare un'osservazione sulla

proposizione a).

In virtù del teor. (2.1) di pag. 138 la continuità dell'operatore lineare A equivale alla continuità nel punto 0, e questa, per il teor. (1.10) di pag. 110, equivale alla continuità "per successioni" nel punto 0, ossia alla proposizione seguente:

a') Per ogni successione $(x_n)_{n\in N}$ di punti di S avente limite 0, risulta:

$$\lim Ax_n = A(0) = 0 .$$

Nella a') si richiede che l'uguaglianza (9) di a) sia verificata per tutte le successioni $(x_n)_{n\in N}$ aventi limite 0, mentre nella a) la (9) si richiede solo per quelle tali che sia convergente la successione $(Ax_n)_{n\in N}$. Dunque

$$a' \Rightarrow a$$

e pertanto (A continuo) $\Rightarrow$ a [in quanto (A continuo) $\Leftrightarrow$ a'].

Orbene, è opportuno notare che senza aver dimostrato il teor. (2.6) non si può dire che, viceversa, a $\Rightarrow$ (A continuo).

— — — — — — — — — — — — — — — — —

Dim. (2.6).

a $\Rightarrow$ b. Per la (1.8) di pag. 109 si tratta di far vedere che, detto G il grafico di A, per ogni successione convergente $(p_n)_{n\in N}$ di punti di G il limite p_0 appartiene a G.

Posto:

$$p_n = (x_n, Ax_n) \quad \forall n \in N$$

$$p_0 = (x_0, y_0)$$

si ha:

$$\lim x_n = x_0 , \quad \lim Ax_n = y_0$$

ossia, posto $x'_n = x_n - x_0$:

$$\lim x'_n = 0 , \quad \lim Ax'_n = \lim (Ax_n - Ax_0) = y_0 - Ax_0 .$$

In virtù della a) ne segue che $y_0 = Ax_0$, ossia che $p_0 \in G$.

— — — — — — — —

b $\Rightarrow$ c. Detto Σ il sostegno di S, definiamo in Σ una nuova norma, ponendo:

$$\|x\|_1 = \|x\| + \|Ax\| \quad \forall x \in \Sigma$$

e denotiamo con S_A lo spazio normato che si ottiene associando tale norma allo spazio vettoriale Σ (nella notazione S_A abbiamo evidenziato la dipendenza della nuova norma dall'operatore A). Gli spazi S, S_A sono due spazi normati aventi lo stesso sostegno Σ: dimostriamo che anche il secondo è uno spazio di Banach.

Denotiamo allora con $(x_n)_{n\in N}$ un'arbitraria successione di Cauchy di S_A.

Dunque per ogni $\varepsilon > 0$ esiste $m \in N$ tale che:

$$\|x_p - x_q\|_1 < \varepsilon \quad \forall p > m \text{ e } \forall q > m ,$$

cioè tale che:

$$\|x_p - x_q\| + \|A(x_p - x_q)\| < \varepsilon \quad \forall p > m \text{ e } \forall q > m .$$

Il primo membro rappresenta anche la distanza, nello spazio $S \times S_1$, delle coppie (x_p, Ax_p), (x_q, Ax_q): dunque la successione

$$\big((x_n, Ax_n) \big)_{n\in N} \tag{10}$$

di punti dello spazio prodotto $S \times S_1$ è una successione di Cauchy. Siccome $S \times S_1$

e' completo in quanto S ed S_1 sono spazi di Banach per ipotesi [cfr. (3.3) pag. 113], la successione (10) e' convergente ; detto (x_0, y_0) il suo limite, si ha allora:

$$\lim \left(\|x_n - x_0\| + \|Ax_n - y_0\| \right) = 0 \,. \tag{11}$$

D'altra parte la (10) e' una successione di punti del grafico G di A, che e' un insieme chiuso, e siccome converge ad (x_0, y_0), il limite deve appartenere a G, ossia deve essere $y_0 = Ax_0$. Di conseguenza la (11) puó scriversi:

$$\lim \|x_n - x_0\|_1 = 0$$

e pertanto la successione $(x_n)_{n \in N}$ e' convergente nello spazio S_A. Resta cosi' provato che S_A e' uno spazio di Banach.

Consideriamo ora l'operatore:

$$I : x \in S_A \to x \in S,$$

che e' un isomorfismo algebrico di S_A su S. Poiche' risulta:

$$\|Ix\|_S = \|x\|_S \leq \|x\|_1 \quad \forall x \in S$$

l'operatore I e' continuo. Per il teorema di Banach [(2.4) pag. 141], che e' applicabile in quanto S_A ed S sono spazi di Banach, ne segue che l'operatore inverso:

$$I^{-1} : x \in S \to x \in S_A$$

e' anch'esso continuo, ossia e' limitato. Esiste allora $\ell > 0$ tale che:

$$\|I^{-1}x\|_{S_A} \leq \ell \|x\|_S \quad \forall x \in S$$

e tale relazione, essendo $I^{-1}x = x$, significa che:

$$\|x\| + \|Ax\| \leq \ell \|x\| \quad \forall x \in S.$$

Se ne deduce la disuguaglianza:

$$\|Ax\| \leq \ell \|x\| \quad \forall x \in S$$

che esprime la continuita' dell'operatore A.

— — — — — — — —

c $\Rightarrow$ a. Gia' stabilita prima della dimostrazione del teorema.

3. Esempi

In questo numero vogliamo far vedere che gli operatori lineari, in particolare i funzionali lineari, considerati negli esempi di pag. 64 sono tutti continui. Aggiungeremo poi qualche ulteriore esempio.

1) Qualunque sia $k \in N$, fissato un punto $(c_1, \dots, c_k) \in R^k$, il funzionale lineare u definito in R^k ponendo:

$$u(x_1, x_2, \dots, x_k) = c_1 x_1 + c_2 x_2 + \dots + c_k x_k \tag{1}$$

e' ovviamente continuo, essendo un polinomio omogeneo di primo grado in k variabili reali. Ricordiamo che tutti i funzionali lineari definiti in R^k sono del tipo (1).

2) Nell'es.2 di pag.64 abbiamo visto che gli operatori lineari di R^k in R^m sono del tipo:

$$x \in R^k \to Ax \in R^m \tag{2}$$

dove A è una matrice di tipo (m,k) ad elementi reali, e il vettore $x=(x_1,\dots,x_k)$ di R^k è identificato con una matrice ad una colonna e k righe. Se si effettua il prodotto di tali matrici e si identifica il vettore colonna ottenuto con la m^{pla} delle sue righe, si vede che le componenti del vettore Ax di R^m sono polinomi di primo grado in k variabili e quindi sono funzioni continue. Pertanto l'operatore (2) è continuo.

3) Detti S_k ed S_m due spazi vettoriali normati a dimensione finita, di dimensioni rispettive k ed m, qualunque operatore lineare di S_k in S_m:

$$A : S_k \to S_m$$

è continuo. Invece se S, S' sono spazi vettoriali normati e il primo di essi non è a dimensione finita, esistono operatori lineari di S in S' che non sono continui: ne indicheremo due esempi dopo aver dimostrato la continuità del predetto A.

Denotato con j_k [con j_m] l'isomorfismo canonico di R^k in S_k [di R^m in S_m] (cfr. (7) pag.141) consideriamo l'operatore B di R^k in R^m che ad ogni $x \in R^k$ associa l'elemento di R^m cui si perviene applicando successivamente [si veda lo schema qui accanto] gli operatori j_k, A, j_m^{-1}; cioè poniamo:

$$\begin{array}{ccc} R^k & \xrightarrow{B} & R^m \\ {\scriptstyle j_k}\downarrow & & \downarrow{\scriptstyle j_m} \\ S_k & \xrightarrow{A} & S_m \end{array}$$

$$B = j_m^{-1} A j_k . \tag{3}$$

L'operatore B, lineare in quanto prodotto di operatori lineari, è continuo perché di R^k in R^m [es. 2)].

Dalla (3), moltiplicando ambo i membri a destra per j_k^{-1} e poi a sinistra per j_m, si deduce l'uguaglianza:

$$A = j_m B j_k^{-1}$$

dalla quale consegue che A è continuo, in quanto prodotto di operatori lineari e continui.

È interessante notare che, se si identifica R^k con S_k ed R^m con S_m [cfr. (2.5) pag.141], gli operatori j_k, j_m divengono identità e l'operatore A di S_k in S_m si identifica con l'operatore B di R^k in R^m che è continuo per l'es. 2).

Denotiamo ora con S il sottospazio vettoriale normato di $C^\circ(R)$ [pag.93 e es.3 pag.100] costituito dalle funzioni continue a supporto compatto in R [si veda la def. dell'es.14 di pag.57, ma non si confonda S con lo spazio vettoriale topologico $C_0^\circ(R)$]. Consideriamo quindi l'operatore lineare L che ad ogni $x \in S$ associa la funzione $t \in R \to tx(t)$, a sua volta elemento di S.

Mostriamo che L non è continuo, ossia [(2.1) pag.138, $c_1 \Leftrightarrow c_5$] che il rapporto $\|Lx\| / \|x\|$ $(x \neq 0)$ non è limitato superiormente. A tale scopo per ogni $n \in N$ consideriamo la funzione x_n (elemento di S) definita in R ponendo $x_n(t)=0$ $\forall t \notin]n-1, n+1[$, $x_n(t)=t-n+1$ per ogni $t \in [n-1, n]$ e $x_n(t)=-t+n+1$ per ogni $t \in]n, n+1]$. La funzione Lx_n, non negativa, è strettamente crescente in $[n-1, n]$ (avendo ivi derivata positiva) e strettamente decrescente in $[n, n+1]$, sicché Lx_n assume il suo massimo valore nel punto n. Si ha allora $\|Lx_n\| = Lx_n(n) = n$, e siccome ovviamente $\|x_n\|=1$, il predetto rapporto nel punto x_n è uguale ad n, sicché non è limitato superiormente.

Con lo stesso ragionamento e con la stessa successione $\{x_n\}$ si prova che il funzionale lineare $M : x \in S \to \int_{-\infty}^{+\infty} tx(t)\,dt$ non è continuo, in quanto risulta $\|Mx_n\| = |Mx_n| = n$ $\forall n \in N$.

4) Fissato $z \in L^2(\Omega)$ (Ω aperto di R^K), il funzionale lineare:

$$u : x \in L^2(\Omega) \longrightarrow \int_\Omega x(t) z(t)\, dt$$

e' continuo.

Infatti, ricordando il prodotto scalare in $L^2(\Omega)$ [es. 2) pag. 99] e tenendo presente la disuguaglianza di Schwarz [cfr. (5) pag. 97] si ha:

$$|u(x)| = |(x,z)| \leq \|x\| \cdot \|z\| \qquad \forall x \in L^2(\Omega)$$

e quindi il funzionale lineare u e' limitato, ossia continuo [cfr. (2.1) pag. 138] (1).

Piu' in generale, se p, p' sono numeri reali >1 tali che $\frac{1}{p} + \frac{1}{p'} = 1$, fissato $z \in L^{p'}(\Omega)$ il funzionale lineare:

$$u : x \in L^p(\Omega) \to \int_\Omega x(t) z(t)\, dt$$

e' continuo.

Infatti, ricordando la disuguaglianza di Schwarz-Hölder [cfr. (3) pag. 54] si ha:

$$|u(x)| \leq \|x\|_{L^p(\Omega)} \cdot \|z\|_{L^{p'}(\Omega)} \qquad \forall x \in L^p(\Omega). \tag{4}$$

5) Fissati $a_{ij} \in C^0(T)$, $b_i \in C^0(T)$, $c \in C^0(T)$ [$i=1,\dots,K$; $j=1,\dots,K$; T pseudodominio di R^K], l'operatore lineare:

$$A : x \in C^2(T) \longrightarrow \sum_{i,j}^{1\dots K} a_{ij} \frac{\partial^2 x}{\partial t_i \partial t_j} + \sum_{i=1}^{K} b_i \frac{\partial x}{\partial t_i} + c x \quad \in C^0(T)$$

e' continuo.

Infatti, posto:

$$\alpha = \max_{i,j} \|a_{ij}\|_{C^0(T)} \quad , \quad \beta = \max_i \|b_i\|_{C^0(T)} \quad , \quad \gamma = \|c\|_{C^0(T)}$$

e successivamente

$$\ell = \max\{\alpha, \beta, \gamma\},$$

per ogni $t \in T$ si ha:

$$|Ax(t)| \leq \ell \left(\sum_{i,j}^{1\dots K} \left| \frac{\partial^2 x(t)}{\partial t_i \partial t_j} \right| + \sum_{i=1}^{K} \left| \frac{\partial x(t)}{\partial t_i} \right| + |x(t)| \right) \leq \ell \left(\sum_{i,j}^{1\dots K} \left\| \frac{\partial^2 x}{\partial t_i \partial t_j} \right\|_{C^0(T)} + \sum_{i=1}^{K} \left\| \frac{\partial x}{\partial t_i} \right\|_{C^0(T)} + \|x\|_{C^0(T)} \right) = \ell \|x\|_{C^2(T)}$$

e quindi:

$$\|Ax\|_{C^0(T)} = \sup_{t \in T} |Ax(t)| \leq \ell \|x\|_{C^2(T)} \qquad \forall x \in C^0(T). \tag{5}$$

Pertanto l'operatore lineare A e' limitato, ossia continuo.

---

(1) Il funzionale u e' un operatore di $L^2(\Omega)$ in R, e la norma di $u(x)$ in R e' il valore assoluto di $u(x)$.

Non si confonda la nozione di "funzionale lineare limitato" [caso particolare di operatore lineare limitato, pag. 137] con quella di "funzione reale limitata" (ossia, dotata di maggioranti e minoranti).

6) Siano p, p' numeri reali tali che $\frac{1}{p}+\frac{1}{p'}=1$, e sia α un multiindice tale che $|\alpha|\le m\in N$. Fissato $z\in W^{m,p'}(\Omega)$, il funzionale lineare:

$$u : x\in W^{m,p}(\Omega) \to \int_\Omega D^\alpha x(t)\, D^\alpha z(t)\, dt$$

è continuo.

Infatti per la disuguaglianza di Schwarz-Hölder [cfr. (3) pag. 54] si ha:

$$|u(x)| \le \|D^\alpha x\|_{L^p(\Omega)} \|D^\alpha z\|_{L^{p'}(\Omega)}$$

e siccome risulta [cfr. es. 13 pag. 103]:

$$(6)\qquad \|D^\alpha x\|_{L^p(\Omega)} = \left(\|D^\alpha x\|^p_{L^p(\Omega)}\right)^{1/p} \le \left(\sum_{0\le|\alpha|\le m} \|D^\alpha x\|^p_{L^p(\Omega)}\right)^{1/p} = \|x\|_{W^{m,p}(\Omega)},$$

nonché, analogamente:

$$\|D^\alpha z\|_{L^{p'}(\Omega)} \le \|z\|_{W^{m,p'}(\Omega)},$$

ne segue:

$$(7)\qquad |u(x)| \le \|x\|_{W^{m,p}(\Omega)} \|z\|_{W^{m,p'}(\Omega)} \qquad \forall x\in W^{m,p}(\Omega),$$

donde la continuità di u.

7) Fissato un multiindice α tale che $|\alpha|\le m\in N$, l'operatore lineare:

$$A : x\in H^m_0(\Omega) \to D^\alpha x \in L^2(\Omega)$$

è continuo.

Infatti per la (6) e per il fatto che la norma in $H^m_0(\Omega)$ coincide con quella in $W^m(\Omega)$, si ha:

$$\|Ax\|_{L^2(\Omega)} = \|D^\alpha x\|_{L^2(\Omega)} \le \|x\|_{H^m_0(\Omega)} \qquad \forall x\in H^m_0(\Omega)$$

e quindi A è continuo.

8) Fissati $a_{ij}\in C^1(\Omega)$ $[i,j=1,\dots,n]$, l'operatore lineare:

$$A : x\in H^2_0(\Omega) \to \sum_{i,j}^{1\dots n} \frac{\partial}{\partial t_i}\left(a_{ij}\frac{\partial x}{\partial t_j}\right) \in L^2(\Omega)$$

è continuo.

Infatti, posto:

$$\alpha = K^2 \max_{i,j} \|a_{ij}\|_{C^1(\Omega)},$$

si ha:

$$\|Ax\|_{L^2(\Omega)} \le \sum_{i,j}^{1\dots n} \left\|\frac{\partial}{\partial t_i}\left(a_{ij}\frac{\partial x}{\partial t_j}\right)\right\|_{L^2(\Omega)} = \sum_{i,j}^{1\dots n} \left\|\frac{\partial a_{ij}}{\partial t_i}\frac{\partial x}{\partial t_j} + a_{ij}\frac{\partial^2 x}{\partial t_j \partial t_i}\right\|_{L^2(\Omega)} \le$$

$$\le \frac{\alpha}{K^2}\sum_{i,j}^{1\dots n}\left[\int_\Omega \left(\left|\frac{\partial x}{\partial t_j}\right| + \left|\frac{\partial^2 x}{\partial t_j \partial t_i}\right|\right)^2 dt\right]^{1/2} \le \frac{2\alpha}{K^2}\sum_{i,j}^{1\dots n}\left[\int_\Omega \left[\left(\frac{\partial x}{\partial t_j}\right)^2 + \left(\frac{\partial^2 x}{\partial t_j \partial t_i}\right)^2\right] dt\right]^{1/2} \le 2\alpha \|x\|_{H^2_1(\Omega)}$$

qualunque sia $x\in H^2_0(\Omega)$. Pertanto risulta:

$$\|Ax\|_{L^2(\Omega)} \leq 2\alpha \|x\|_{H_0^2(\Omega)} \qquad \forall x \in H_0^2(\Omega)$$

donde la continuità di A.

9) Procedendo come nell'esempio precedente, si verifica che fissati $a_{ij} \in C^0(\Omega)$ $[i,j=1,\dots,n]$ l'operatore lineare:

$$A : x \in H_0^2(\Omega) \longrightarrow \sum_{i,j}^{1\dots n} a_{ij} \frac{\partial^2 x}{\partial t_i \partial t_j} \in L^2(\Omega)$$

è continuo.

10) Fissato $z \in L^2(\Omega)$, il funzionale lineare:

$$u : x \in H_0^1(\Omega) \longrightarrow \int_\Omega x(t) z(t)\, dt$$

è continuo.

Si ha infatti:

$$|u(x)| \leq \|x\|_{L^2(\Omega)} \|z\|_{L^2(\Omega)} \leq \|z\|_{L^2(\Omega)} \|x\|_{H_0^1(\Omega)} \qquad \forall x \in H_0^1(\Omega).$$

11) Fissati $q_0, q_1, q_2 \in L^2(\Omega)$ con Ω aperto di R, il funzionale lineare:

$$u : x \in H_0^2(\Omega) \longrightarrow \int_\Omega (q_0 x + q_1 x' + q_2 x'')\, dt$$

è continuo.

Infatti per ogni $x \in H_0^2(\Omega)$ si ha:

$$|u(x)| \leq \int_\Omega |q_0| \cdot |x|\, dt + \int_\Omega |q_1| \cdot |x'|\, dt + \int_\Omega |q_2| \cdot |x''|\, dt \leq$$

$$\leq \|q_0\|_{L^2(\Omega)} \|x\|_{L^2(\Omega)} + \|q_1\|_{L^2(\Omega)} \|x'\|_{L^2(\Omega)} + \|q_2\|_{L^1(\Omega)} \|x''\|_{L^2(\Omega)}$$

$$\leq \left(\|q_0\|_{L^2(\Omega)} + \|q_1\|_{L^2(\Omega)} + \|q_2\|_{L^2(\Omega)} \right) \left(\|x\|_{L^2(\Omega)} + \|x'\|_{L^2(\Omega)} + \|x''\|_{L^2(\Omega)} \right)$$

ossia, detta l la somma delle norme in $L^2(\Omega)$ di q_0, q_1, q_2, risulta:

$$|u(x)| \leq l \|x\|_{H_0^2(\Omega)} \qquad \forall x \in H_0^2(\Omega).$$

12) Fissati $z \in L^2(\Omega)$, $E \in L^\infty(\Omega)$ con $\Omega \subseteq R$, nel seguito incontreremo il funzionale lineare:

$$u : x \in H \longrightarrow \int_\Omega z\,(x - Ex'')\, dt$$

dove $H = H_0^1(\Omega) \cap H^2(\Omega)$, la norma in H essendo quella di $H^2(\Omega)$.

Mostriamo che u è continuo.

Per ogni $x \in H$ si ha:

$$|u(x)| \le \int_\Omega |u||z|\,dt + \int_\Omega |z||E||x'|\,dt \le \|z\|_{L^2(\Omega)}\|x\|_{L^2(\Omega)} + \sup|E|\|z\|_{L^2(\Omega)}\|x'\|_{L^2(\Omega)} \le$$
$$\le \|z\|_{L^2(\Omega)}\|x\|_{L^2(\Omega)} + \sup|E|\|z\|_{L^2(\Omega)}\|x'\|_{L^2(\Omega)} + \|z\|_{L^2(\Omega)}\|x''\|_{L^2(\Omega)} \le$$
$$\le \|z\|_{L^2(\Omega)}(1+\sup|E|)(\|x\|_{L^2(\Omega)} + \|x'\|_{L^2(\Omega)} + \|x''\|_{L^2(\Omega)})$$

e quindi, detto h il prodotto dei primi due fattori all'ultimo membro, risulta:

$$|u(x)| \le h\|x\|_H \qquad \forall x \in H\,.$$

13) *Distribuzioni su un aperto* Ω.

Come esempio di funzionali lineari e continui, riporteremo tra breve quelli che si chiamano *distribuzioni su un aperto,* e introdurremo qualche nozione relativa ad esse. Con l'occasione, ci soffermeremo sull'importante *distribuzione di Dirac.*

Preliminarmente è opportuno effettuare alcuni richiami su quanto esposto nei numeri 6 e 7 del capitolo II (pag.44-45), con qualche aggiunta utile per il seguito.

Detta $\mathcal{L}_\Omega$ la totalità dei sottoinsiemi di Ω misurabili secondo Lebesgue, per ogni punto $z \ge 0$ dello spazio $L^1_{loc}(\Omega)$ [cfr.n.17 pag.58] la corrispondenza

$$\mu : X \in \mathcal{L}_\Omega \to \int_X z(t)\,dt \tag{8}$$

è una *misura in* $\mathcal{L}_\Omega$, che assume in X il valore $+\infty$ se ivi la funzione z non è sommabile; la funzione z che interviene nella (8) si chiama **la densità** di μ, e perciò si dice che la (8) è una **misura dotata di densità**. Si dimostra che per f funzione misurabile in Ω ed $X \subseteq \Omega$ misurabile secondo Lebesgue, sussiste l'equivalenza:

(9) *f sommabile in X rispetto a μ $\Leftrightarrow$ fz sommabile in X secondo Lebesgue* ;

inoltre, se f è sommabile in X rispetto a μ, *l'integrale di Stieltjes-Lebesgue rispetto a μ della funzione f coincide con l'integrale secondo Lebesgue della funzione fz* :

$$\int_X f\,d\mu = \int_X f(t)\,z(t)\,dt \qquad \forall X \in \mathcal{L}_\Omega\,. \tag{10}$$

Vedremo tra breve che esistono misure non dotate di densità. Per una misura μ del tipo (8) e per $X \in \mathcal{L}_\Omega$ sussistono banalmente le implicazioni:

$$X \text{ compatto} \Rightarrow \mu(X) < +\infty \tag{11}$$

$$X \text{ di misura nulla secondo Lebesgue} \Rightarrow \mu(X) = 0\,, \tag{12}$$

e viceversa, per il teorema di **Nikodym**, qualunque misura μ in $\mathcal{L}_\Omega$ che verifichi tali implicazioni è dotata di densità ; ossia, per ogni siffatta μ esiste una funzione $z \ge 0$ appartenente ad $L^1_{loc}(\Omega)$ che consente di esprimere μ mediante la (8).

Vogliamo ora soffermarci su un importante esempio di misura non dotata di densità.

Indichiamo con $\mathfrak{L}^{(k)}$ la totalità dei sottoinsiemi di $\mathbf{R}^k$ misurabili secondo Lebesgue e fissiamo un punto $t_0 \in \mathbf{R}^k$: è facile constatare che la corrispondenza

$$\mu_{t_0} : X \in \mathfrak{L}^{(k)} \to \begin{cases} 1 & \text{se } t_0 \in X \\ 0 & \text{se } t_0 \notin X \end{cases} \tag{13}$$

è una misura in $\mathfrak{L}^{(k)}$; mostriamo che essa è priva di densità. Infatti, se μ_{t_0} fosse uguale alla funzione μ definita dalla (8) [con $\Omega = \mathbf{R}^k$ e quindi $\mathfrak{L}_\Omega = \mathfrak{L}_{\mathbf{R}^k} = \mathfrak{L}^{(k)}$] si avrebbe $\mu_{t_0}(t_0) = 0$ (perché l'integrale (8) con $X = \{t_0\}$ è nullo, in quanto $\{t_0\}$ ha misura nulla secondo Lebesgue), e ciò è assurdo giacché per la (13) risulta $\mu_{t_0}(t_0) = 1$.

Dimostriamo che per la misura (13) sussiste la proposizione:

I. *Ogni funzione f reale o complessa, misurabile in* $\mathbf{R}^k$ *e continua in un intorno del punto* t_0, *è sommabile rispetto a* μ_{t_0} *in tutto lo spazio* $\mathbf{R}^k$ *e risulta* :

$$\int_{\mathbf{R}^k} f \, d\mu_{t_0} = f(t_0) \quad \forall t_0 \in \mathbf{R}^k \ , \tag{14}$$

e che per k=1 si ha pure :

II. *Ogni funzione reale o complessa f di una variabile reale, misurabile in* $[t_0, +\infty[$ *e continua in un intorno destro del punto* t_0, *è sommabile rispetto a* μ_{t_0} *in* $[t_0, +\infty[$, *e risulta*:

$$\int_{[t_0, +\infty[} f \, d\mu_{t_0} = f(t_0) \quad \forall t_0 \in \mathbf{R} \ (^1). \tag{15}$$

DIM.(14). Per ogni $n \in \mathbf{N}$ sia Γ_n il cerchio chiuso di $\mathbf{R}^k$ di centro t_0 e raggio $1/n$: siccome per la (13) risulta $\mu_{t_0}(\mathbf{R}^k - \Gamma_n) = 0$, la funzione f è sommabile rispetto a μ_{t_0} nell'insieme $\mathbf{R}^k - \Gamma_n$, e l'integrale di f rispetto a μ_{t_0} esteso a tale insieme è nullo. Scelto $\nu \in \mathbf{N}$ tale che in Γ_ν la f sia continua, per ogni $n \geq \nu$ la f è sommabile rispetto a μ_{t_0} anche in Γ_n ; pertanto f è sommabile rispetto a μ_{t_0} in tutto $\mathbf{R}^k$, e risulta

$$\int_{\mathbf{R}^k} f \, d\mu_{t_0} = \int_{\Gamma_n} f \, d\mu_{t_0} \quad \forall n \geq \nu \, . \tag{16}$$

D'altra parte, per il teorema della media, per ogni $n \geq \nu$ esiste un punto $t_{0,n} \in \Gamma_n$ tale che l'integrale a secondo membro della (16) è uguale ad $f(t_{0,n})\mu_{t_0}(\Gamma_n)$, ossia ad $f(t_{0,n})$ dato che $\mu_{t_0}(\Gamma_n) = 1$ [cfr.(13)]. Dunque per la (16) si ha :

$$\int_{\mathbf{R}^k} f \, d\mu_{t_0} = f(t_{0,n}) \quad \forall n \geq \nu \, . \tag{17}$$

---

(1) Il lettore osservi che si ha

$$\int_{]t_0, +\infty[} f \, d\mu_{t_0} = 0 \, . \tag{*}$$

Infatti , siccome t_0 non appartiene all'insieme di integrazione, ivi la μ_{t_0} si annulla e quindi l'integrale è nullo. Dunque la relazione (*) è diversa dalla (15) se $f(t_0) \neq 0$.

Ciò evidenzia che, contrariamente a quanto accade per l'integrale secondo Lebesgue, non è lecito scrivere il primo membro della (15) nella forma

$$\int_{t_0}^{+\infty} f(t) \, d\mu_{t_0}$$

dalla quale non emerge se t_0 appartiene o non all'insieme di integrazione.

Ne consegue che la successione di termine generale $f(t_{0,n})$ è definitivamente costante, sicchè $f(t_{0,n}) = \lim f(t_{0,n})$; d'altra parte quest'ultimo è uguale ad $f(t_0)$, perché $t_{0,n} \to t_0$ (in quanto $|t_{0,n} - t_0| < 1/n$) e la funzione f è continua in t_0. Dunque $f(t_{0,n}) = f(t_0) \quad \forall n \geq \nu$: perciò la (17) coincide con la (14) e l'asserto è provato.

DIM.(15). Il ragionamento appena concluso è valido anche ponendo in esso k=1 e $\Gamma_n = [t_0, t_0 + 1/n]$.

* * *

Ciò premesso, fissiamo un aperto Ω di $\mathbf{R}^k$ (anche k=1) . I funzionali lineari definiti e continui nello spazio $C_0^\infty(\Omega)$ [es.16 pag.58 ed es.11 pag.102] si chiamano **distribuzioni su** Ω [perciò lo spazio $C_0^\infty(\Omega)$ si suole denotare anche con $\mathcal{D}(\Omega)$. Si dice che una distribuzione T è **positiva,** e si scrive $T \geq 0$, quando per $\varphi \in C_0^\infty(\Omega)$ si ha:

(18) $$\varphi \geq 0 \;\Rightarrow T(\varphi) \geq 0$$ ([2]).

Siccome $C_0^\infty(\Omega)$ non è uno spazio normato, la continuità di un funzionale lineare $u(x)$ definito per $x \in C_0^\infty(\Omega)$ non si può stabilire, come negli esempi precedenti, mediante un'opportuna maggiorazione di $|u(x)|$. Poiché $C_0^\infty(\Omega)$ non è neppure uno spazio metrizzabile, non si può applicare il teorema (1.10) di pag.110 per dedurne che la continuità di u equivale alla continuità "*per successioni*"(cfr. fine pag.109). Nel citato teorema, peraltro, l'ipotesi "spazio metrico"non è necessaria per la validità della tesi: basta un'ipotesi più generale (che qui non interessa formulare); ma neppure tale ipotesi è verificata dallo spazio $C_0^\infty(\Omega)$. Tuttavia è stato dimostrato che:

(19) *La continuità di un funzionale lineare definito in* $C_0^\infty(\Omega)$ *equivale alla continuità per successioni*.

Esempi di distribuzioni.

1. Distribuzioni "di tipo T_z **"**

Fissato un punto z dello spazio $L^1_{loc}(\Omega)$, mostriamo che il funzionale lineare:

(20) $$T_z : \varphi \in C_0^\infty(\Omega) \to \int_\Omega z(t)\varphi(t)\,dt$$

è continuo.

([2]) Da evidenziare che nella notazione $T \geq 0$, e quindi nella (18), si usa il termine *positiva* nonostante il segno $\geq$.

A chiarimento della (18), il lettore osservi che, essendo T lineare, la scrittura $T \geq 0$ non può significare (in conformità con la stessa notazione adoperata per esprimere che una funzione non assume valori negativi)

(*) $$T(\varphi) \geq 0 \quad \forall \varphi \in C_0^\infty(\Omega),$$

in quanto la (*) comporta $T(\varphi) = 0 \;\forall \varphi \in C_0^\infty(\Omega)$. Infatti per ogni $\varphi \in C_0^\infty(\Omega)$, essendo anche $-\varphi \in C_0^\infty(\Omega)$, dalla (*) consegue $T(-\varphi) \geq 0$, ossia per la linearità $-T(\varphi) \geq 0$ e quindi $T(\varphi) \leq 0$: se ne deduce che $T(\varphi) = 0$.

DIM. Per la (19), si tratta di stabilire per ogni fissato $\varphi \in C_0^\infty(\Omega)$ quanto segue: se $\{\varphi_n\}_{n\in\mathbf{N}}$ è una qualunque successione convergente di punti di $C_0^\infty(\Omega)$ tale che

(21) $$\lim \varphi_n = \varphi \quad \text{in } C_0^\infty(\Omega),$$

risulta:

(22) $$\lim T_z(\varphi_n) = T_z(\varphi) \ .$$

Ricordando il significato della (21) [cfr. pag.102, es.11) e poi es.3)] , esiste un compatto $K \subset \Omega$ che contiene i supporti di φ e di $\varphi_n \ \forall n \in \mathbf{N}$, tale che si abbia, uniformemente in K e qualunque sia il multiindice α :

(23) $$\lim_{n\to\infty} \varphi_n = \varphi \quad , \quad \lim_{n\to\infty} D^\alpha \varphi_n = D^\alpha \varphi .$$

Fissato $\varepsilon > 0$ e posto

(24) $$h = 1 + \int_K |z(t)|\, dt \ ,$$

per la prima delle (23) esiste $\nu \in \mathbf{N}$ tale che si abbia:

$$|\varphi_n(t) - \varphi(t)| < \frac{\varepsilon}{h} \quad \forall n > \nu \ \text{ e } \ \forall t \in K .$$

Dunque per ogni $n > \nu$, tenendo presente che in $\Omega - K$ le φ_n, φ sono identicamente nulle, risulta:

$$|T_z(\varphi_n) - T_z(\varphi)| \le \int_\Omega |z(t)||\varphi_n(t) - \varphi(t)|\, dt = \int_K |z(t)||\varphi_n(t) - \varphi(t)|\, dt < \frac{\varepsilon}{h}\int_K |z(t)|\, dt < \varepsilon$$

donde la (22).

Dunque il funzionale (20) è una distribuzione su Ω .

Tenuto conto anche del fatto che sussistono le equivalenze:

(25) $$T_z \ge 0 \Leftrightarrow z \ge 0 \quad , \quad T_z = 0 \Leftrightarrow z = 0 \ ,$$

delle quali non riportiamo la dimostrazione, spesso la funzione $z \in L^1_{loc}(\Omega)$ viene identificata con la distribuzione T_z , e ciò si può esprimere dicendo che la *z* viene intesa ***nel senso delle distribuzioni***, con un linguaggio analogo a quello che nel seguito esprimerà una importante nozione concernente le funzioni appartenenti ad $L^1_{loc}(\Omega)$.

Nota.

Tenendo presente che la (19) è vera anche se si sostituisce $C_0^\infty(\Omega)$ con $C_0^0(\Omega)$ [anche questo è stato dimostrato], la dimostrazione dianzi esposta prova che anche il funzionale definito nello spazio $C_0^0(\Omega)$ [cfr. es.14 pag.57 e es. 9) pag.102]

(26) $$M_z : \varphi \in C_0^0(\Omega) \to \int_\Omega z(t)\varphi(t)\, dt$$

è continuo: basta osservare che delle due relazioni (23) è stata utilizzata solo la prima. I funzionali lineari definiti in $C_0^0(\Omega)$ ed ivi continui si possono identificare con le misure in $\mathcal{L}_\Omega$ soddisfacenti a certe condizioni (che qui non è il caso di riportare) : perciò tali funzionali si chiamano **misure su** Ω .

Con tale linguaggio, il prolungamento per continuità su $C_0^0(\Omega)$ di una distribuzione del tipo T_z è una misura su Ω .

2. Distribuzioni "di tipo T_μ".

Fissata una misura μ in $\mathcal{L}_\Omega$ per la quale sussista l'implicazione (11) di pag.149, il funzionale lineare [cfr.n.7 pag.45]

(27) $$T_\mu : \varphi \in C_0^\infty(\Omega) \to \int_\Omega \varphi \; d\mu \qquad (^3)$$

è continuo : la dimostrazione è del tutto analoga a quella dianzi effettuata per il funzionale T_z (^{4}). Dunque T_μ è una distribuzione su Ω.

Come accade per le distribuzioni di tipo T_z , anche il prolungamento per continuità su $C_0^0(\Omega)$ di una distribuzione di tipo T_μ è una misura su Ω.

Quando si vuole esprimere che una distribuzione T è di tipo T_μ , si suol dire che T ***è un integrale*** [dato che $T_\mu(\varphi)$ si esprime mediante un integrale]. Anche $T_z(\varphi)$ si esprime mediante un integrale, ma quando $T = T_z$ si preferisce evidenziare che T_z si identifica con la funzione z , e perciò in tal caso si suol dire che T ***è una funzione***. [sottinteso ***appartenente ad*** $L^1_{loc}(\Omega)$].

Ovviamente le distribuzioni di tipo T_μ sono positive [cfr.(18) pag.151]; viceversa sussiste il teorema:

(28) *Ogni distribuzione positiva è di tipo T_μ : cioè, se T è una distribuzione positiva, esiste una misura μ in L_Ω che assume valore finito negli insiemi compatti, tale che:*

$$T(\varphi) = \int_\Omega \varphi \; d\mu \quad \forall \varphi \in C_0^\infty(\Omega).$$

Mostriamo che esistono distribuzioni simultaneamente di tipo T_μ e di tipo T_z .

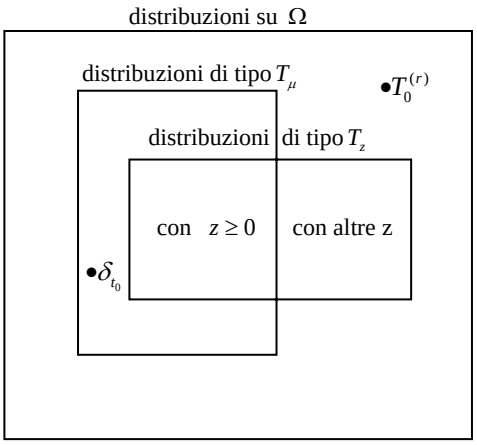
Infatti , fissata una distribuzione T_μ , se per la misura μ sussistono le (11), (12) di pag.149 , per il teorema di Nikodym [pag.149] esiste una funzione $z \geq 0$, appartenente ad $L^1_{loc}(\Omega)$, tale che μ coincide con la (8) . Allora [cfr. (10) con $X = \Omega$ ed $f = \varphi$] si ha $T_\mu(\varphi) = T_z(\varphi) \;\; \forall \varphi \in C_0^\infty(\Omega)$, ossia $T_\mu = T_z$.

Se una distribuzione T_μ è anche di tipo T_z , necessariamente risulta $z \geq 0$ [essendo di tipo T_μ si tratta di una distribuzione positiva , e siccome è anche di tipo T_z , da $T_z \geq 0$ per la (25) segue $z \geq 0$] ; una distribuzione di tipo T_z con $z \geq 0$ è ovviamente positiva, e quindi per la (28) è anche di tipo T_μ .

Tali conclusioni sono illustrate nello schema della pagina seguente.

(^{3}) Per il significato dell'integrale osserviamo quanto segue. Per ogni $\varphi \in C_0^\infty(\Omega)$, detto S_φ il supporto di φ , la φ è identicamente nulla in $\Omega - S_\varphi$, quindi ivi sommabile rispetto alla misura μ , ed è tale anche nell'insieme compatto S_φ [cfr.(16) pag.150 con z costante uguale ad 1] : pertanto rispetto a μ la φ è sommabile in Ω .

(^{4}) Formalmente la dimostrazione si deduce da quella citata, con riferimento a T_μ , sostituendo negli integrali che in essa figurano $|z(t)|dt$ con $d\mu$ nonché $\varphi_n(t), \varphi(t)$ rispettivamente con φ_n, φ ; inoltre nella (24) l'integrale va sostituito con $\mu(K)$ [qui interviene l'ipotesi (14) che μ assuma valore finito negli insiemi compatti].



Le distribuzioni δ_{t_0} e $T_0^{(r)}$ indicate nello schema saranno introdotte rispettivamente nell'es. **3** (qui di seguito) e nell'es. **4** [di pag.157].

3. **La distribuzione di Dirac.**

Fissato un punto $t_0 \in \mathbf{R}^k$, il funzionale lineare

(29) $$\delta_{t_0} : \varphi \in C_0^\infty(\mathbf{R}^k) \to \varphi(t_0)$$

è ovviamente continuo ([5]), e quindi è una distribuzione su $\mathbf{R}^k$.

La (29) si chiama **la distribuzione di Dirac**, **relativa al punto** t_0 ; peraltro spesso viene chiamata **la** δ **di Dirac**, per un motivo che sarà chiaro tra breve.

Notiamo che , per la (14) di pag.150 con $f = \varphi$, dalla (29) consegue:

(30) $$\delta_{t_0}(\varphi) = \varphi(t_0) = \int_{\mathbf{R}^k} \varphi \; d\mu_{t_0} \; .$$

Abbiamo dianzi osservato che esistono distribuzioni che sono simultaneamente di tipo T_μ e di tipo T_z. Orbene, dimostriamo che [si veda lo schema qui sopra]:

La distribuzione di Dirac δ_{t_0} *è di tipo* T_μ*, ma non è di tipo* T_z

ossia, con il linguaggio di pag.153, *<u>la distribuzione di Dirac è un integrale, ma non è una funzione</u>*.

DIM. Fissato $t_0 \in \mathbf{R}^k$, denotando per semplicità con T_0 la distribuzione T_μ con $\mu = \mu_{t_0}$ e $\Omega = \mathbf{R}^k$ [cfr. (27) pag.152] dalla (30) si deduce,

$$\delta_{t_0}(\varphi) = T_0(\varphi) \qquad \forall \varphi \in C_0^\infty(\mathbf{R}^k) \, ,$$

sicché $\delta_{t_0} = T_0$: pertanto δ_{t_0} è di tipo T_μ.

Per dimostrare che δ_{t_0} non è una funzione, procedendo per assurdo ammettiamo che esista $z \in L^1_{loc}(\mathbf{R}^k)$ tale che δ_{t_0} sia la distribuzione T_z [cfr.(20)pag.151 con $\Omega = \mathbf{R}^k$], sicché risulta [cfr.(30)]:

([5]) Dalla (21) di pag.151, il cui significato è stato richiamato dopo la (22), segue banalmente che

$$\lim_{n\to\infty} \varphi_n(t_0) = \varphi(t_0) \quad \text{ossia} \quad \lim_{n\to\infty} \delta_{t_0}(\varphi_n) = \delta_{t_0}(\varphi) \, .$$

(31) $$\varphi(t_0) = \int_{\mathbf{R}^k} z(t)\,\varphi(t)\,dt \quad \forall \varphi \in C_0^\infty(\mathbf{R}^k)\,.$$

Poiché ovviamente $C_0^\infty(\mathbf{R}^k - \{t_0\}) \subset C_0^\infty(\mathbf{R}^k)$ ([6]), la (31) si può applicare con $\varphi = \psi \in C_0^\infty(\mathbf{R}^k - \{t_0\})$, e sostituendo l'insieme di integrazione con $\mathbf{R}^k - \{t_0\}$ [l'integrale non cambia] ; siccome $\psi(t_0) = 0$, se ne deduce che

$$\int_{\mathbf{R}^k - \{t_0\}} z(t)\,\psi(t)\,dt = 0 \quad \forall \psi \in C_0^\infty(\mathbf{R}^k - \{t_0\})\,.$$

Dunque la distribuzione su $\mathbf{R}^k - \{t_0\}$ determinata da *z* è identicamente nulla, e da ciò per la seconda delle (25) [pag.152]segue che la funzione $z \in L^1_{loc}(\mathbf{R}^k)$ di cui abbiamo ammesso l'esistenza è uguale a 0 quasi ovunque in $\mathbf{R}^k - \{t_0\}$, quindi anche quasi ovunque in $\mathbf{R}^k$: allora l'integrale nella (31) è nullo, e perciò risulta

$$\varphi(t_0) = 0 \quad \forall \varphi \in C_0^\infty(\mathbf{R}^k)\ ,$$

il che è assurdo ([7]).

Riportiamo qui l'uguaglianza che nella (30) evidenzia due espressioni di $\delta_{t_0}(\varphi)$ e che, per ricordare al lettore che questa è la distribuzione di Dirac calcolata nel punto φ, indicheremo con il simbolo (D) anziché con un numero:

(D) $$\int_{\mathbf{R}^k} \varphi\, d\mu_{t_0} = \varphi(t_0) \qquad \forall \varphi \in C_0^\infty(\mathbf{R}^k)\,.$$

Orbene, è ben noto che la (D) interviene frequentemente nelle scienze applicate, ma abitualmente si scrive nella forma

(D') $$\int_{\mathbf{R}^k} \varphi(t)\delta(t - t_0)\,dt = \varphi(t_0) \qquad \forall \varphi \in C_0^\infty(\mathbf{R}^k)\,,$$

nella quale la funzione δ, definita ponendo

(32) $$\delta(t) = 0 \quad \forall t \in \mathbf{R}^k - \{0\} \quad , \quad \delta(0) = +\infty\,,$$

([6]) Una funzione $\psi \in C_0^\infty(\mathbf{R}^k - \{t_0\})$, che ha il supposto contenuto in $\mathbf{R}^k - \{t_0\}$ ed è ivi di classe C^∞, risulta di classe C^∞ in tutto $\mathbf{R}^k$ dato che in $\mathbf{R}^k - \{t_0\}$ è identicamente nulla, e quindi appartiene anche a $C_0^\infty(\mathbf{R}^k)$.

([7]) Per completezza occorre provare che esiste almeno una funzione $\varphi \in C_0^\infty(\mathbf{R}^k)$ che non si annulla in un dato punto t_0 ; è sufficiente stabilire ciò per *k*=1 e t_0=0. A tale scopo consideriamo la funzione φ definita ponendo:

(*) $$\varphi(t) = e^{\frac{1}{t^2-1}} \quad \forall t \in]-1,1[\quad , \quad \varphi(t) = 0 \text{ per } |t| \geq 1\,,$$

che è a supporto compatto in **R e** non si annulla nel punto 0. Mostriamo che $\varphi \in C_0^\infty(\mathbf{R}^k)$.

Per $t \to \pm 1$ la φ è infinitesima, quindi è continua nei due punti; la derivata di φ *di qualunque ordine* in un punto $t \in]-1,1[$ è una somma , e in ciascuno degli addendi figura sempre il fattore $e^{\frac{1}{t^2-1}}$, il quale per $t \to \pm 1$ è un infinitesimo di ordine infinitamente grande; quindi la derivata *di qualunque ordine* tende a zero per $t \to \pm 1$: ne segue che la φ è indefinitamente derivabile nei punti ± 1, e che la derivata di qualunque ordine è ivi nulla: dopo ciò è chiaro che $\varphi \in C_0^\infty(\mathbf{R}^k)$ e l'asserto è provato.

è tale da risultare

(33) $$\int_{\mathbf{R}^k} \delta(t)dt = 1 \ .$$

Diciamo subito che, essendo la funzione δ nulla quasi ovunque in $\mathbf{R}^k$ come risulta dalla (32), l' integrale al primo membro della (D') , ossia

(34) $$\int_{\mathbf{R}^k} \varphi(t)\delta(t-t_0)dt \,,$$

è uguale a 0, e così pure quello al primo membro della (33) : perciò l'uguaglianza (33) è falsa, e così pure è falsa la (D') per tutte le funzioni $\varphi \in C_0^\infty(\mathbf{R}^k)$ che non si annullano nel punto t_0.

Dal confronto di (D), (D') emerge che la (D') si adopera come se fosse vera la relazione

(35) $$\int_{\mathbf{R}^k} \varphi \, d\mu_{t_0} = \int_{\mathbf{R}^k} \varphi(t)\delta(t-t_0)\, dt \quad \forall \varphi \in C_0^\infty(\mathbf{R}^k) \,,$$

che esprime l'uguaglianza nel punto φ di due distribuzioni : la distribuzione di Dirac T_μ con $\mu = \mu_{t_0}$ e la distribuzione T_z con $z(t) = \delta(t-t_0)$; in altri termini, si utilizza la (35) come se la misura μ_{t_0} fosse dotata di densità. Ma la (35) è assurda perché la distribuzione di Dirac è di tipo T_μ ma non è di tipo T_z [come abbiamo dimostrato a pag.154]; in altri termini, perché la misura μ_{t_0} non è dotata di densità [come abbiamo dimostrato a pag. 150].

A questo punto vien da chiedersi come mai una relazione falsa, appunto la (D'), sia comunemente adoperata rivelandosi utilissima in svariate questioni, anche di settori diversi l'uno dall'altro. La risposta è molto semplice: l'integrale (34) che figura nella (D') va inteso non come un integrale secondo Lebesgue (cioè interpretando il simbolo correttamente), ma come una notazione comodissima nella pratica (anche se abusiva) per indicare il primo membro della (D), ossia l'integrale di Stieltjes-Lebesgue

(36) $$\int_{\mathbf{R}^k} \varphi \, d\mu_{t_0} \,,$$

nel quale la misura è quella definita dalla (13) di pag .150.

Analogamente l'integrale

(37) $$\int_{\mathbf{R}^k} \delta(t)\, dt$$

si deve intendere come un modo diverso (sia pure abusivo) di scrivere l'integrale

$$\int_{\mathbf{R}^k} d\mu_{t_0} \,,$$

che effettivamente è uguale a $\mu_{t_0}(\mathbf{R}^k)$, ossia ad 1 stante la citata definizione di μ_{t_0} .

La scrittura (34) spiega l'origine dell'uso della lettera δ per indicare la distribuzione di Dirac ([8]), quando non occorre precisare il punto t_0. Probabilmente la scelta della lettera δ deriva dall'iniziale del cognome Dirac.

Rileviamo infine che anche i primi membri delle due relazioni seguenti, frequentemente adoperate:

(38) $$\int_{-\infty}^{+\infty} e^{-2\pi j\omega t}\delta(t)\,dt = 1 \quad \forall \omega \in \mathbf{R} \quad , \quad \int_{-\infty}^{+\infty} e^{-st}\delta(t)\,dt = 1 \quad \forall s \in \mathbf{C}$$

vanno intesi non come integrali secondo Lebesgue (sarebbero uguali a zero), ma come notazioni improprie per indicare rispettivamente gli integrali

(39) $$\int_{\mathbf{R}} f_\omega \, d\mu_0 \quad , \quad \int_{[0,+\infty[} g_s \, d\mu_0$$

dove , per la precisione dei due simboli, abbiamo indicato la funzione integranda ponendo rispettivamente per ogni $\omega \in \mathbf{R}$ e per ogni $s \in \mathbf{C}$:

$$f_\omega(t) = e^{-2\pi j\omega t} \quad \forall t \in \mathbf{R} \quad , \quad g_s(t) = e^{-st} \quad \forall t \in \mathbf{R}\,.$$

Naturalmente nella (39) μ_0 è la misura (13) di pag.150 con t_0=0 [nel primo integrale anche con k=1].

Per la dimostrazione delle due relazioni (38), intese come appena detto, basta applicare con t_0=0 le proposizioni enunciate a pag. 150 e poi dimostrate: infatti la (14) con $f = f_\omega$ e k=1 è la prima delle (38) , la (15) con $g = g_s$ è la seconda; naturalmente occorre tener presente che $f_\omega(0) = g_s(0) = 1$.

Le (38) esprimono rispettivamente **la trasformata di Fourier** e **la trasformata di Laplace** della δ di Dirac ([9]) : è interessante notare che ciò non emerge se i due integrali si scrivono correttamente nella forma (39).

Nota. A conclusione dei discorsi sulla distribuzione di Dirac, vogliamo accennare alla presunta origine sia della funzione δ definita dalla (32) di pag.155, sia della notazione (34) e dell'uguaglianza (33) di pag156, supponendo per semplicità k=1 e t_0=0.

La comunità scientifica aveva bisogno di una funzione $\delta(t)$ definita in **R** tale da verificare la relazione:

(40) $$\int_{-\infty}^{+\infty} \varphi(t)\delta(t)\,dt = \varphi(0) \qquad \forall \varphi \in C_0^\infty(\mathbf{R})\,,$$

ossia la (D') di pag.155 scritta per k=1 e t_0=0; inoltre la $\delta(t)$ doveva verificare la (33).

Si è posto allora il problema di ricercare una siffatta funzione δ , considerando la (40) come equazione nell'incognita $\delta(t)$: esponiamo il procedimento che presumibilmente è stato seguito.

Detta $u = u(t)$ la funzione scalino relativa al punto 0, qualunque sia $\varphi \in C_0^\infty(\mathbf{R})$ si ha [per semplicità, scriveremo spesso la funzione integranda ritenendo sottintesa la variabile t]:

$$\int_{-\infty}^{+\infty} u\varphi'\,dt = \int_{-\infty}^{0} u\varphi'\,dt + \int_{0}^{+\infty} u\varphi'\,dt = \int_{0}^{+\infty} \varphi'\,dt = \left[\varphi(t)\right]_0^{+\infty} = \lim_{t\to\infty}\varphi(t) - \varphi(0) = -\varphi(0)$$

([8]) Notiamo che per $t_0 = 0$ la scrittura (34) formalmente significa $T_\delta(\varphi)$ [cfr.(20) pag.151 con $z = \delta$, $t_0 = 0$, e $\Omega = \mathbf{R}^k$], e ricordiamo che la T_δ [che in effetti è la distribuzione di Dirac relativa al punto 0, pur non essendo essa di tipo T_z] si suole identificare con la funzione δ [pag.152 dopo la (25)] .

([9]) Rinunciamo ad introdurre la trasformata di Fourier [di Laplace] " di una distribuzione su Ω ", che è un certo funzionale lineare e continuo, ma non è una distribuzione perché non è definito in tutto $C_0^\infty(\Omega)$.

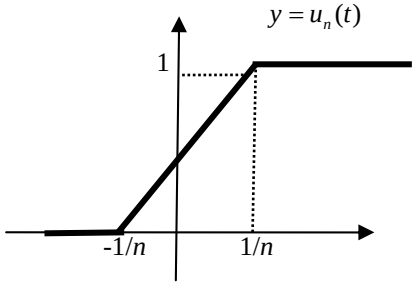
e quindi risulta:

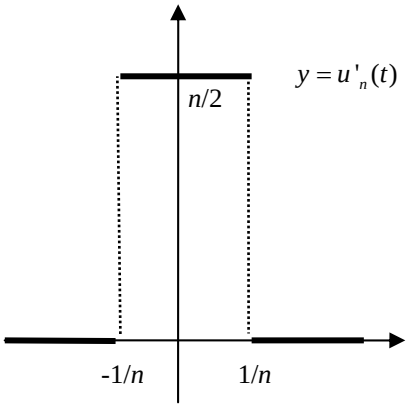
$$\varphi(0) = -\int_{-\infty}^{+\infty} u\varphi'\,dt \,. \tag{41}$$

Pertanto la (40) , intesa come equazione nell'incognita $\delta(t)$, si può scrivere:

$$\int_{-\infty}^{+\infty} \varphi(t)\delta(t)\,dt = -\int_{-\infty}^{+\infty} u(t)\varphi'(t)\,dt \,. \tag{42}$$

Osserviamo che la funzione scalino *u* relativa al punto 0 può essere approssimata puntualmente in $\mathbf{R} - \{0\}$ mediante i termini di una successione $\{u_n\}$ di funzioni continue. Una siffatta successione si ottiene ponendo [si veda il primo dei diagrammi seguenti]:





$$u_n(t) = \begin{cases} 0 \;\; se \;\; t < -1/n \\ \dfrac{1}{2}(nt+1) \;\; se \;\; t \in [-\dfrac{1}{n}, \dfrac{1}{n}] \\ 1 \;\; se \;\; t > 1/n \end{cases} \tag{43}$$

sicché risulta:

$$\lim_{n\to\infty} u_n(t) = u(t) \quad \forall t \in \mathbf{R} - \{0\} \,.$$

Per trasformare opportunamente il secondo membro della (42) consideriamo le uguaglianze seguenti, che non sono tutte lecite e perciò scriveremo il simbolo $\overset{?}{=}$ (sovrapponendo un punto interrogativo al segno =) dove l'uguaglianza non è assicurata:

$$-\int_{-\infty}^{+\infty} u\varphi'\,dt = -\int_{-\infty}^{+\infty} (\lim_{n\to\infty} u_n)\,\varphi'\,dt \overset{?}{=} -\lim_{n\to\infty} \int_{-\infty}^{+\infty} u_n\varphi'\,dt = -\lim_{n\to\infty} [u_n\varphi]_{-\infty}^{+\infty} + \lim_{n\to\infty} \int_{-\infty}^{+\infty} u'_n\,\varphi\;\,dt =$$

$$= \lim_{n\to\infty} \int_{-\infty}^{+\infty} u'_n\,\varphi\;\,dt \overset{?}{=} \int_{-\infty}^{+\infty} (\lim_{n\to\infty} u'_n)\,\varphi\;\,dt \,.$$

Sostituendo il secondo membro della (42) con l'ultimo di tali integrali, si conclude che : se l'equazione (42) ammette una soluzione $\delta(t)$, e se sono valide le uguaglianze indicate con $\overset{?}{=}$, deve risultare

$$\int_{-\infty}^{+\infty} \varphi(t)\,\delta(t)\,dt = \int_{-\infty}^{+\infty} (\lim_{n\to\infty} u'_n(t))\,\varphi(t)\,dt \qquad \forall \varphi \in C_0^\infty(\mathbf{R})$$

ossia

$$\int_{-\infty}^{+\infty} [\delta(t) - \lim_{n\to\infty} u'_n(t)]\,\varphi(t)\,dt = 0 \quad \forall \varphi \in C_0^\infty(\mathbf{R}) \,.$$

Orbene, tale relazione sussiste se e solo se risulta:

$$\delta(t) = \lim_{n\to\infty} u'_n(t) \quad \text{q.o. in } \mathbf{R} \quad (^{10}), \tag{44}$$

e siccome per la (43) si ha:

(^{10}) Ciò si deduce dalla seconda delle (25) di pag.152, utilizzata tenendo conto della (20).

$$u'_n(t)=\begin{cases}0 \ \ se \ \ t<-1/n\\ \dfrac{n}{2} \ \ se \ \ t\in[-\dfrac{1}{n},\dfrac{1}{n}]\\ 0 \ \ se \ \ t>1/n\end{cases}$$

dalla (44) consegue che la soluzione $\delta(t)$ della (42), se esiste, deve essere la funzione δ definita dalla (32) di pag.155. A questo punto occorrerebbe controllare se tale funzione δ verifica la (42), ed è ragionevole presumere che ciò non sia stato fatto : perchè con la predetta funzione la (42), equivalente alla (40) ossia alla (D') con $t_0=0$, non è soddisfatta [abbiamo visto a pag.155 che la (D') è falsa].

Quanto alla (33) [abbiamo osservato che è falsa, e chiarito come va intesa], si può presumere che sia stato seguito il seguente procedimento. Siccome l'integrale

$$\int_{-\infty}^{+\infty} u'_n(t)\,dt$$

è uguale [si veda il secondo diagramma di pag.157] all'area a_n del rettangolo $[-\dfrac{1}{n},\dfrac{1}{n}]\times[0,\dfrac{n}{2}]$, ossia ad 1 qualunque sia $n\in\mathbf{N}$, si ha:

$$\int_{-\infty}^{+\infty}\delta(t)\,dt=\int_{-\infty}^{+\infty}[\lim_{n\to\infty}u'_n(t)]\,dt\overset{?}{=}\lim_{n\to\infty}\int_{-\infty}^{+\infty}u'_n(t)\,dt=\lim_{n\to\infty}a_n=1\,.$$

Ma l'uguaglianza indicata con $\overset{?}{=}$ non è valida, altrimenti sarebbero uguali il primo e l'ultimo membro delle precedenti relazioni il che, come già detto, è falso.

4. **Ulteriori esempi di distribuzioni.**

Abbiamo visto le distribuzioni "di tipo T_z", quelle "di tipo T_μ", quelle che sono simultaneamente di entrambi i tipi, e la distribuzione di Dirac che è di tipo T_μ ma non di tipo T_z .

Vogliamo mostrare che esistono distribuzioni che non sono "di tipo T_z" e neppure "di tipo T_μ". Infatti , detto Ω un aperto di **R** , qualunque sia $r\in\mathbf{N}$ il funzionale lineare:

$$T_0^{(r)}:\varphi\in C_0^\infty(\Omega)\to\varphi^{(r)}(0)$$

è continuo ([11]), e quindi è una distribuzione su Ω .

([11]) Si tratta di stabilire per ogni $\varphi\in C_0^\infty(\Omega)$ quanto segue: se $\{\varphi_n\}_{n\in\mathbf{N}}$ è una qualunque successione convergente di punti di $C_0^\infty(\Omega)$ tale che

(*) $$\lim\varphi_n=\varphi \quad \text{in } C_0^\infty(\Omega),$$

risulta:

(**) $$\lim\varphi_n^{(r)}(0)=\varphi^{(r)}(0)\ .$$

Ricordando il significato della (*) [cfr. pag.102, es.11) e poi es.3] , esiste un compatto $K\subset\Omega$ che contiene i supporti di φ e di $\varphi_n\ \forall n\in\mathbf{N}$, tale che si abbia, uniformemente in K e qualunque sia $r\in\mathbf{N}$:

$$\lim_{n\to\infty}\varphi_n=\varphi \quad , \quad \lim_{n\to\infty}\varphi_n^{(r)}=\varphi^{(r)}\,.$$

Dalla seconda di tali relazioni segue banalmente la (**) se $0\in K$; se $0\notin K$ la (**) è banale , dato che i numeri che figurano nei due menbri sono uguali a 0.

Ovviamente la $T_0^{(r)}$ non è una distribuzione positiva, e quindi non è di tipo T_μ [cfr.(28) pag.153]. Per dimostrare che la $T_0^{(r)}$ non è di tipo T_z si può ripetere il ragionamento esposto nella DIM. di pag.154 per provare che non è di tipo T_z la distribuzione di Dirac δ_{t_0} : basta porre ivi t_0=0 e sostituire T_0 con $T_0^{(r)}, \varphi$ con $\varphi^{(r)}$; nella nota ([5]) di pag.154 si sostituiranno le parole "*che non si annulla*" con "*la cui derivata di ordine r non si annulla*".

Derivate di una distribuzione. Derivate nel senso delle distribuzioni.

Qualunque sia la distribuzione *T* su un aperto $\Omega \subseteq \mathbf{R}^k$, e qualunque sia il multiindice α, si chiama **derivata di** *T* **di multiindice** α, e si denota con $D^\alpha T$, il funzionale lineare definito come segue:

(45) $$D^\alpha T : \varphi \in C_0^\infty(\Omega) \to (-1)^{|\alpha|} T(D^\alpha \varphi)$$

che è continuo ([12]) e quindi è a sua volta una distribuzione.

Dunque per calcolare la derivata $D^\alpha T$ nel punto φ , formalmente basta sostituire nell'espressione di $T(\varphi)$ la funzione φ con la sua derivata $D^\alpha \varphi$ e moltiplicare per $(-1)^{|\alpha|}$.

Ad esempio per $k = 4$ ed $\alpha = (3,2,0,5)$, sicché $|\alpha| = 10$, qualunque sia la distribuzione *T* si ha:

$$D^{(3,2,0,5)} T = T\left(\frac{\partial^{10} \varphi}{\partial t_1^3 \partial t_2^2 \partial t_4^5}\right) .$$

La nozione introdotta consente di esprimere facilmente la seguente: qualunque sia la funzione $z \in L^1_{loc}(\Omega)$, la derivata $D^\alpha T_z$ si chiama **derivata di** *z* di multiindice α **nel senso delle distribuzioni**, e si denota col solito simbolo $D^\alpha z$, : dunque quest'ultima si ottiene identificando la funzione *z* con la distribuzione T_z. Tenendo presente che la distribuzione T_z è [cfr. (20) pag.151]

$$T_z : \psi \in C_0^\infty(\Omega) \to \int_\Omega z(t)\, \psi(t)\, dt ,$$

per la (45) con $T = T_z$ si ha allora:

(46) $$(D^\alpha z)(\varphi) = (D^\alpha T_z)(\varphi) = (-1)^{|\alpha|} \int_\Omega z(t)(D^\alpha \varphi)(t)\, dt \quad \forall \varphi \in C_0^\infty(\Omega) .$$

Ad esempio si ha, per ogni $i \in \{1, ..., k\}$:

(47) $$\frac{\partial^2 z}{\partial t_i^2}(\varphi) = \int_\Omega z(t) \frac{\partial^2 \varphi}{\partial t_i^2}(t) dt \qquad \forall \varphi \in C_0^\infty(\Omega) .$$

([12]) Ovviamente se $\varphi_n \to \varphi$ in $C_0^\infty(\Omega)$ [cfr. pag.102, es.11) e poi es.3), tenendo presente quanto detto a pag.101 prima dell'esempio 6)] , in tale spazio si ha anche $D^\alpha \varphi_n \to D^\alpha \varphi$: dopo ciò per la continuità di *T* si ha $T(D^\alpha \varphi_n) \to T(D^\alpha \varphi)$, e quindi [cfr.(45)] $D^\alpha T(\varphi_n) \to D^\alpha T(\varphi)$.

Non è detto che la distribuzione $D^\alpha z$, derivata nel senso delle distribuzioni della funzione z, *sia a sua volta una funzione* (si veda la locuzione introdotta a pag.153), cioè non è detto che esista una funzione $w \in L^1_{loc}(\Omega)$ tale che si abbia:

(48) $$\int_\Omega w(t)\,\varphi(t)\,dt = (-1)^{|\alpha|}\int_\Omega z(t)(D^\alpha\varphi)(t)\,dt \quad \forall \varphi \in C_0^\infty(\Omega).$$

Ricordiamo [cfr.(5) pag.61] che se (e solo se) esiste una siffatta w, la funzione $z \in L^1_{loc}(\Omega)$ è dotata della derivata debole $D^\alpha z$ e per definizione si ha $w = D^\alpha z$: dunque se z è dotata della derivata debole $D^\alpha z$, la derivata di z nel senso delle distribuzioni si identifica con la derivata debole di z (elemento di $L^1_{loc}(\Omega)$), e perciò in tal caso [e solo in tal caso] *è una funzione*.

Il lettore osservi che qualunque elemento dello spazio $L^1_{loc}(\Omega)$ ammette nel senso delle distribuzioni qualunque derivata, mentre non è detto che sia dotato della stessa derivata in senso debole; qualora ne sia dotato, le due derivate si identificano [peraltro, anche se si denotano con lo stesso simbolo, non si può dire che *coincidono*, perché una di esse è una distribuzione, l'altra è un elemento di $L^1_{loc}(\Omega)$].

Ad esempio la funzione scalino $u = u(t - t_0)$ relativa al punto t_0, derivabile quasi ovunque in $\mathbf{R}$ con derivata quasi ovunque nulla, non è dotata di derivata debole : prima di stabilire ciò, calcoliamo la derivata (di ordine 1) di $u(t - t_0)$ nel senso delle distribuzioni, ossia la derivata della distribuzione:

$$T_u : \varphi \in C_0^\infty(\Omega) \to \int_{-\infty}^{+\infty} u(t - t_0)\,\varphi(t)\,dt \quad .$$

Per la (45) con $\alpha = 1$ si ha, tenendo presente la (41):

$$DT_u(\varphi) = -T_u(\varphi') = -\int_{-\infty}^{+\infty} u(t - t_0)\,\varphi'(t)\,dt = \varphi(t_0) \quad \forall \varphi \in C_0^\infty(\Omega)$$

e quindi, per la (29) pag.154 :

$$DT_u(\varphi) = \delta_{t_0}(\varphi) \quad \forall \varphi \in C_0^\infty(\Omega)$$

ossia $DT_u = \delta_{t_0}$. Abbiamo così stabilito che:

La derivata nel senso delle distribuzioni della funzione scalino, relativa al punto t_0, è la distribuzione di Dirac δ_{t_0}. Equivalentemente: la funzione scalino è una primitiva (*nel senso delle distribuzioni*) *della distribuzione δ di Dirac.*

La seconda formulazione ha indotto a denotare la funzione scalino relativa al punto 0 col simbolo δ_{-1} .

Dopo ciò è chiaro che la funzione scalino non è dotata della derivata debole u': se ne fosse dotata, la derivata nel senso delle distribuzioni, ossia la distribuzione di Dirac δ_{t_0}, dovrebbe identificarsi con un elemento di $L^1_{loc}(\mathbf{R}^k)$ [pag.160 dopo la (48)], cioè dovrebbe essere di tipo T_z e ciò, come abbiamo visto, non accade

Per esprimere la derivata nel senso delle distribuzioni di una funzione $z \in L^1_{loc}(\Omega)$, abbiamo specificato la derivata di una distribuzione di tipo T_z. A titolo di esempio, vogliamo ora specificare le derivate di distribuzioni di altro tipo.

La derivata di multiindice α di una distribuzione di tipo T_μ [(27) pag.152] è la distribuzione:

$$D^\alpha T_\mu : \varphi \in C_0^\infty(\Omega) \to (-1)^{|\alpha|} \int_\Omega D^\alpha \varphi \, d\mu ,$$

e quella della distribuzione di Dirac δ_{t_0} relativa al punto $t_0 \in \mathbf{R}^k$ [(29) pag.154] è la distribuzione definita ponendo:

$$(49) \qquad (D^\alpha \delta_{t_0})(\varphi) = (-1)^{|\alpha|} (D^\alpha \varphi)(t_0) = (-1)^{|\alpha|} \int_{\mathbf{R}^k} (D^\alpha \varphi) \, d\mu_{t_0} \ .$$

Ricordando che l'integrale nella (49) abitualmente si scrive in forma diversa [cfr.(D),(D') pag.155 , con $D^\alpha \varphi$ in luogo di φ] , la (49) di solito si scrive nella forma seguente:

$$(D^\alpha \delta_{t_0})(\varphi) = (-1)^{|\alpha|} \int_{\mathbf{R}^k} (D^\alpha \varphi)(t) \, \delta(t - t_0) \, dt \ ;$$

in particolare per $k = 1$ e $t_0 = 0$, quindi con $\alpha \in \mathbf{N}$ (preferiamo allora indicarlo con n) tale relazione si scrive:

$$(D^n \delta)(\varphi) = (-1)^n \int_{-\infty}^{+\infty} \varphi^{(n)}(t) \, \delta(t) \, dt \ .$$

Per $\Omega \subseteq \mathbf{R}$ e $t_0 = 0$, qualunque siano $r, n \in \mathbf{N}$ la derivata di ordine n della distribuzione $T_0^{(r)}$ [cfr.**4.** pag.158] è la distribuzione:

$$D^n T_0^{(r)} : \varphi \in C_0^\infty(\Omega) \to (-1)^n (D^n \varphi^{(r)})(0)$$

e quindi

$$D^n T_0^{(r)}(\varphi) = (-1)^n \varphi^{(r+n)}(0) \ .$$

A proposito della derivata di una distribuzione, ci sembra utile segnalare il seguente teorema ([13]):

Per ogni funzionale lineare e continuo F definito in $C^\infty(\Omega)$, esiste una distribuzione di tipo T_z primitiva di multiindice α di F : precisamente, esistono un multiindice α ed una funzione $z \in L^1_{loc}(\Omega)$ tali che $F = D^\alpha T_z$ in $C_0^\infty(\Omega)$

Concludiamo i cenni sulla distribuzione di Dirac riportando nella forma solitamente adoperata le relazioni (14),(15) di pag. 150 :

$$(50) \qquad \int_{\mathbf{R}^k} f(t)\delta(t - t_0) \, dt = f(t_0) \quad \forall t_0 \in \mathbf{R}^k \ , \qquad \int_{t_0}^{+\infty} f(t)\delta(t - t_0) \, dt = f(t_0) \quad \forall t_0 \in \mathbf{R} \ ,$$

valide la prima qualunque sia f funzione reale o complessa, misurabile in $\mathbf{R}^k$ e continua in un intorno del punto t_0, la seconda qualunque sia f funzione reale o complessa di una variabile reale, misurabile in $[t_0, +\infty[$ e continua in un intorno destro del punto t_0 .

([13]) Cfr. ad es. Kirillov-Gvishiani, *Theorems and problems in Functional Analysis*, 33 pag.83.

Naturalmente per le (50) vale quanto si è detto per la (D') di pag.154: i primi membri vanno intesi non nel senso di Lebesgue (altrimenti sarebbero nulli) ma rispettivamente come i seguenti integrali di Stieltjes-Lebesgue [pag.45].

$$\int_{\mathbf{R}^k} f \, d\mu_{t_0} \quad , \quad \int_{[t_0,+\infty[} f \, d\mu_{t_0}$$

rispetto alla misura (13) di pag.150. Il lettore osservi che questi ultimi si ottengono formalmente dagli integrali (50) sostituendo $\delta(t-t_0)\,dt$ con $d\mu_{t_0}$ e scrivendo f invece di $f(t)$.

* * *

Abbiamo visto a pag.60 che nel caso $\Omega \subseteq \mathbf{R}$ un'equazione differenziale si può intendere *in senso debole*. Naturalmente ciò può farsi anche se $\Omega \subseteq \mathbf{R}^k$, ad es. per l'equazione

(51) $$D^\alpha u = f \ ,$$

che **in senso debole** esprime il problema seguente: assegnata una funzione $f \in L^1_{loc}(\Omega)$, ricercare le funzioni $u \in L^1_{loc}(\Omega)$ dotate della derivata debole $D^\alpha u$ che verificano la (51). Una soluzione di tale problema si chiama **soluzione debole** della (51).

Orbene, la nozione di derivata nel senso delle distribuzioni consente una ulteriore generalizzazione del problema in senso classico: assegnata una distribuzione f su Ω, ricercare le funzioni $u \in L^1_{loc}(\Omega)$ soddisfacenti alla (51), nella quale la derivata è intesa nel senso delle distribuzioni. Allora, denotando per semplicità con <v,*x*> il valore del funzionale v nel punto *x*, cioè ponendo v(*x*)= <v,*x*> ([14]), la (51) significa che:

$$< D^\alpha u, \varphi > \ = \ < f, \varphi > \qquad \forall \varphi \in C_0^\infty(\Omega)$$

cioè che:

(52) $$(-1)^{|\alpha|} \int_\Omega u(t) D^\alpha \varphi(t)\, dt \ = < f, \varphi > \qquad \forall \varphi \in C_0^\infty(\Omega) \, .$$

Se la distribuzione f è *una funzione*, ossia è identificata con una funzione $f \in L^1_{loc}(\Omega)$ [in particolare $f \in L^P(\Omega)$, cfr. pag.59 primo comma], la (52) diviene:

(53) $$(-1)^{|\alpha|} \int_\Omega u(t) D^\alpha \varphi(t)\, dt = \int_\Omega f(t) \varphi(t)\, dt \qquad \forall \varphi \in C_0^\infty(\Omega) \, .$$

La (53) significa che u è dotata della derivata debole $D^\alpha u$, e che l'elemento f di $L^1_{loc}(\Omega)$ è la derivata debole $D^\alpha u$. Pertanto nell'ipotesi anzidetta una soluzione del problema posto è una soluzione debole della (51).

([14]) Di tale notazione ci occuperemo nel seguito [cfr.(1) pag.171].

4. Completamento di uno spazio vettoriale normato, e di uno spazio metrico.

In questo numero dimostreremo il seguente teorema:

(4.1) Ogni spazio vettoriale normato S è equivalente ad un sottospazio di uno spazio di Banach $\hat{S}$, ovunque denso in $\hat{S}$.
Lo spazio $\hat{S}$ è univocamente determinato a meno di isomorfismi isometrici.

Se si conviene di identificare ogni spazio vettoriale normato Σ con tutti gli spazi vettoriali normati ad esso equivalenti, il teorema (4.1) si può enunciare come segue: per ogni spazio vettoriale normato S, esiste uno ed un solo spazio di Banach $\hat{S}$ del quale S sia un sottospazio ovunque denso. Tale asserzione si può esprimere dicendo che ogni spazio vettoriale normato S può essere "ampliato", univocamente, in modo da ottenere uno spazio completo $\hat{S}$ nel quale S è ovunque denso. Il fatto che S è ovunque denso in $\hat{S}$, cioè che la chiusura di S coincide con $\hat{S}$, si potrebbe esprimere con linguaggio intuitivo dicendo che $\hat{S}$ è "il più piccolo" spazio completo di cui S è sottospazio.

Quanto sopra ha lo scopo di rendere naturale la seguente definizione.

Per ogni spazio vettoriale normato S, si chiama completamento di S uno spazio di Banach $\hat{S}$, la cui esistenza è assicurata dal teor. (4.1), godente della seguente proprietà: esiste un sottospazio di $\hat{S}$ equivalente ad S e denso in $\hat{S}$.

Adottando la convenzione sopra precisata, un qualunque spazio di Banach $\hat{S}$ che abbia la predetta proprietà si chiama il completamento di S.

La nozione di completamento si introduce anche con riferimento ad uno spazio metrico, in quanto sussiste anche il teorema seguente:

(4.2). Ad ogni spazio metrico S si può associare uno spazio metrico completo $\hat{S}$, univocamente determinato a meno di isometrie, godente della seguente proprietà: esiste un'isometria di S su un sottospazio metrico S_1 di $\hat{S}$ ovunque denso in $\hat{S}$.

Convenendo di identificare ogni spazio metrico Σ con qualunque spazio metrico Σ_1 tale che esista un'isometria di Σ su Σ_1, lo spazio metrico completo $\hat{S}$ menzionato in (4.2) si chiama il completamento dello spazio metrico S.

Ci limiteremo a dimostrare il teorema (4.1), relativo ad uno spazio vettoriale normato. Peraltro, salvo alcune modifiche prevalentemente formali, il procedimento dimostrativo del teorema (4.2) è analogo a quello del (4.1).

Dim. (4.1). Per chiarezza di esposizione, procederemo attraverso le seguenti tappe:

a) Costruzione, a partire da S, di un opportuno spazio vettoriale $\hat{S}$.

b) Introduzione di una norma in $\hat{S}$.

c) Completezza dello spazio vettoriale normato $\hat{S}$.

d) Esistenza di un sottospazio S_1 di $\hat{S}$ equivalente ad S.

e) Il sottospazio S_1 è ovunque denso in $\hat{S}$.

f) Ogni spazio di Banach Σ godente delle proprietà di $\hat{S}$ è equivalente ad $\hat{S}$.

---

a) Sia $\mathcal{C}$ l'insieme costituito dalle successioni di Cauchy di punti di S [1]. Introduciamo in $\mathcal{C}$ l'usuale operazione di addizione, denotando, qualunque siano $c'=(x'_n)_{n\in N}\in\mathcal{C}$ e $c''=(x''_n)_{n\in N}\in\mathcal{C}$, con $c'+c''$ la successione $(x'_n+x''_n)_{n\in N}$, che è ovviamente un elemento di $\mathcal{C}$; introduciamo poi la moltiplicazione per uno scalare, denotando, per ogni $\lambda\in K$ e per ogni $c=(x_n)_{n\in N}\in\mathcal{C}$, con λc la successione $(\lambda x_n)_{n\in N}\in\mathcal{C}$. Con tali operazioni, la terna $(\mathcal{C},+,\cdot)$ è uno spazio vettoriale, che indicheremo semplicemente con $\mathcal{C}$.

Sia ora H il sottospazio vettoriale di $\mathcal{C}$ costituito dalle successioni $(x_n)_{n\in N}$ di punti di S convergenti al punto 0 di S, e sia $\hat{S}$ lo spazio vettoriale quoziente $\mathcal{C}/H$.

Pertanto [cfr. pag. 53] gli elementi di $\hat{S}$ sono le classi di equivalenza modulo H degli elementi di $\mathcal{C}$; se $(x_n)_{n\in N}$ è un elemento di $\mathcal{C}$, la classe di equivalenza di $(x_n)_{n\in N}$ modulo H è l'insieme X costituito da tutti gli elementi $(y_n)_{n\in N}$ di $\mathcal{C}$ tali che $\lim(x_n-y_n)=0$.

---

b) Sia $X\in\hat{S}$. Per ogni rappresentante $(x_n)_{n\in N}$ della classe X, risultando:

$$(*)\qquad \big|\|x_p\|-\|x_q\|\big|\le\|x_p-x_q\| \qquad \forall p,q\in N$$

la successione delle norme $(\|x_n\|)_{n\in N}$ è convergente, perché è una successione di Cauchy di numeri reali. D'altra parte, se $(y_n)_{n\in N}$ è un altro rappresentante della classe X, sicché $(x_n-y_n)\to 0$, risultando:

$$\big|\|x_n\|-\|y_n\|\big|\le\|x_n-y_n\| \qquad \forall n\in N,$$

---

[1] Naturalmente il teorema è banale se lo spazio vettoriale normato S è completo (basta assumere $\hat{S}=S$). In caso contrario, gli elementi di $\mathcal{C}$ non sono tutte le successioni convergenti.

si ha:

$$\lim_{n\to\omega} \|x_n\| = \lim_{n\to\omega} \|y_n\| .$$

Pertanto é lecito porre:

$$\nu(X) = \lim_{n\to\omega} \|x_n\| .$$

Abbiamo così definito un'applicazione ν di $\hat{S}$ in $[0,+\infty[$ che, come facilmente si verifica, è una norma per $\hat{S}$. Lo spazio vettoriale normato $(\hat{S},\nu)$ sarà denotato semplicemente con $\hat{S}$.

---

c) Premettiamo una osservazione. Siano: $(y_k)_{k\in N}$ un elemento di $\mathcal{C}$, e Y l'elemento di $\hat{S}$ rappresentato da $(y_k)_{k\in N}$. Per ogni $k\in N$, sia Y_k l'elemento di $\hat{S}$ rappresentato dalla successione costante $y_k, y_k, \dots, y_k, \dots$ la quale é un elemento di $\mathcal{C}$ in quanto convergente: noi diremo che $(Y_k)_{k\in N}$ é la successione di elementi di $\hat{S}$ <u>generata</u> da $(y_k)_{k\in N}$. Orbene, é immediato verificare che:

$$(*) \qquad \lim_{k\to\omega} Y_k = Y$$

Infatti, essendo $(y_n)_{n\in N}$ una successione di Cauchy, per ogni $\varepsilon > 0$ esiste $\nu\in N$ tale che:

$$\|y_n - y_k\| < \varepsilon/2 \quad \forall n>\nu \ e \ \forall k>\nu$$

e quindi tale che:

$$\lim_{n\to\omega} \|y_n - y_k\| \le \varepsilon/2 < \varepsilon \quad \forall k>\nu .$$

Poiché il primo membro di tale relazione é la norma in $\hat{S}$ di $Y - Y_k$, la $(*)$ é acquisita.

Ciò premesso, dimostriamo la c).
Sia $(X_k)_{k\in N}$ una successione di Cauchy di elementi di $\hat{S}$: si tratta di dimostrare che essa é convergente . Per ogni $k\in N$, sia $(x_{k,n})_{n\in N}$ un rappresentante di X_k; poiché $(x_{k,n})_{n\in N}$ é una successione di Cauchy, esiste $\tau_k \in N$ tale da risul<u></u>tare:

$$(**) \qquad \|x_{k,n} - x_{k,\tau_k}\| < \frac{1}{k} \quad \forall n \ge \tau_k .$$

La successione $(x_{k,\tau_k})_{k\in N}$ così costruita é anch'essa una successione di Cauchy . Infatti, qualunque siano $k,h\in N$ risulta, per ciascun $n\in N$ tale che $n\ge\tau_k$, $n\ge\tau_h$ [cfr. $(**)$] :

$$\|x_{k,\tau_k} - x_{h,\tau_h}\| \le \|x_{k,\tau_k} - x_{k,n}\| + \|x_{k,n} - x_{h,n}\| + \|x_{h,n} - x_{h,\tau_h}\| <$$

$$< \frac{1}{k} + \|x_{k,n} - x_{h,n}\| + \frac{1}{h}$$

e quindi, passando al limite per $n\to\omega$:

$$\| x_{k,n_k} - x_{h,n_h} \| < \frac{1}{k} + \| X_k - X_h \| + \frac{1}{h} .$$

Da tale relazione consegue l'asserzione fatta, in quanto $(X_k)_{k\in N}$ è una successione di Cauchy di $\hat{S}$.

Detta $(Y_k)_{k\in N}$ la successione di elementi di $\hat{S}$ generata da $(x_{k,n_k})_{k\in N}$, e detto Y l'elemento di $\hat{S}$ rappresentato da $(x_{k,n_k})_{k\in N}$, in virtù della premessa fatta all'inizio risulta [cfr. (*)] :

$$\lim_{k\to\infty} Y_k = Y$$

e d'altro canto la (**) implica, passando al limite per $n \to \infty$:

$$\| X_k - Y_k \| \leq \frac{1}{k} \quad \forall k \in N .$$

Dopo ciò, dalla relazione :

$$\| X_k - Y \| \leq \| X_k - Y_k \| + \| Y_k - Y \| \qquad \forall k \in N$$

consegue che la successione $(X_k)_{k\in N}$ è convergente ed ha per limite Y.

---

d) Per ogni $x \in S$ denotiamo con Ax l'elemento di $\hat{S}$ rappresentato dalla successione costante $x, x, \ldots, x, \ldots\ldots$. Abbiamo così definito un operatore A di S in $\hat{S}$, ovviamente lineare. Posto $S_1 = A(S)$, S_1 è un sottospazio di $\hat{S}$ e l'operatore A è un isomorfismo isometrico di S su S_1 [cfr. pag. 140], in quanto risulta:

$$\| Ax \|_{S_1} = \| Ax \|_{\hat{S}} = \lim_{n\to\infty} \| x \|_S = \| x \|_S \qquad \forall x \in S .$$

Pertanto lo spazio S è equivalente al sottospazio S_1 di $\hat{S}$.

---

e) Sia X un elemento di $\hat{S}$, e sia $\varepsilon > 0$. Detto $(x_n)_{n\in N}$ un rappresentante di X , esiste $k_\varepsilon \in N$ tale che:

$$\| x_m - x_{k_\varepsilon} \| < \varepsilon/2 \qquad \forall m \geq k_\varepsilon$$

in quanto $(x_n)_{n\in N}$ è una successione di Cauchy . Tale relazione implica:

$$\lim_{n\to\infty} \| x_n - x_{k_\varepsilon} \| \leq \varepsilon/2 < \varepsilon ,$$

ossia, detto X_ε l'elemento di $\hat{S}$ rappresentato dalla successione costante $x_{k_\varepsilon}, x_{k_\varepsilon}, \ldots\ldots$:

$$\| X - X_\varepsilon \| < \varepsilon .$$

Pertanto a ciascun intorno di X appartiene almeno un elemento di S_1 , e quindi X è aderente ad S_1 .

Quanto sopra dimostra che la chiusura di S_1 è $\hat{S}$.

---

f) Sia Σ uno spazio di Banach godente delle proprietà di $\hat{S}$: esistono allora un sottospazio Σ_1 di Σ ovunque denso in Σ ,e un isomorfismo isometrico B di S su Σ_1 .

Per ogni $y \in \Sigma$, sia $(y_n)_{n\in N}$ una successione di elementi di Σ_1 convergente (in Σ) ad y ; posto $x_n = B^{-1}y_n \ \forall n\in N$, ovviamente $(x_n)_{n\in N}$ é una successione di Cauchy di S, giacché qualunque siano $p,q \in N$ risulta:

$$\|x_p - x_q\| = \|B^{-1}y_p - B^{-1}y_q\| = \|B^{-1}(y_p - y_q)\| = \|y_p - y_q\| .$$

Sia X l'elemento di $\hat{S}$ rappresentato da $(x_n)_{n\in N}$: evidentemente X non dipende dalla successione $(y_n)_{n\in N}$ di elementi di Σ_1 avente per limite y : infatti, se $(y'_n)_{n\in N}$ é un'altra successione di elementi di Σ_1 avente per limite y , posto $x'_n = B^{-1}y'_n \ \forall n\in N$, risulta:

$$\|x_n - x'_n\| = \|B^{-1}y_n - B^{-1}y'_n\| = \|B^{-1}(y_n - y'_n)\| = \|y_n - y'_n\|$$

e quindi $\lim_{n\to\infty}(x_n - x'_n) = 0$; pertanto l'elemento di $\hat{S}$ rappresentato da $(x'_n)_{n\in N}$ coincide con X .

In definitiva, l'elemento X di $\hat{S}$ ottenuto nel modo precisato dipende solo dall'elemento y di Σ : poniamo $X = Ly$.

Resta così definito un operatore L di Σ in $\hat{S}$ che, come facilmente si verifica,é lineare.

Dimostriamo che $L(\Sigma) = \hat{S}$. All'uopo, sia X un elemento di $\hat{S}$, e sia $(x_n)_{n\in N}$ un rappresentante di X. Posto $y_n = Bx_n \ \forall n\in N$, $(y_n)_{n\in N}$ é una successione di Cauchy di Σ , in quanto risulta, qualunque siano $p,q \in N$:

$$\|y_p - y_q\| = \|Bx_p - Bx_q\| = \|B(x_p - x_q)\| = \|x_p - x_q\| .$$

Poiché Σ é completo, esiste il limite di $(y_n)_{n\in N}$, diciamolo y . Ovviamente risulta $X = Ly$, e quindi $X \in L(\Sigma)$.

Essendo $L(\Sigma) = \hat{S}$, e risultando, per ogni $y \in \hat{\Sigma}$ (con ovvio significato delle notazioni):

$$\|Ly\| = \|X\| = \lim_{n\to\infty}\|x_n\| = \lim_{n\to\infty}\|B^{-1}y_n\| = \lim_{n\to\infty}\|y_n\| = \|y\| \quad (*)$$

si conclude che L é un isomorfismo isometrico di Σ su $\hat{S}$.

(*) Se una successione $(y_n)_{n\in N}$ di punti di uno spazio vettoriale normato converge ed ha limite y, la successione delle norme $(\|y_n\|)_{n\in N}$ converge ed ha per limite $\|y\|$, in quanto risulta:

$$|\|y_n\| - \|y\|| \le \|y_n - y\| \quad \forall n\in N.$$

Esempio.

Consideriamo lo spazio vettoriale normato non completo introdotto alla fine del n.3 del Cap.V [cfr. pag. 115], che abbiamo indicato con $\mathbb{C}^{m,p}(\Omega)$.

Detto A l'operatore che ad ogni $x \in \mathbb{C}^{m,p}(\Omega)$ associa l'elemento di $W^{m,p}(\Omega)$ costituito da tutte le funzioni q.o. uguali ad x, la norma di x in $\mathbb{C}^{m,p}(\Omega)$ è stata definita come la norma di Ax in $W^{m,p}(\Omega)$, sicchè risulta:

$$\|Ax\|_{W^{m,p}(\Omega)} = \|x\|_{\mathbb{C}^{m,p}(\Omega)} .$$

Se indichiamo con H il sottospazio di $W^{m,p}(\Omega)$ codominio dell'operatore A (1), ne segue che A è un isomorfismo isometrico di $\mathbb{C}^{m,p}(\Omega)$ su H. Siccome la chiusura di H è lo spazio normato completo $H^{m,p}(\Omega)$, si conclude che $H^{m,p}(\Omega)$ è il completamento dello spazio $\mathbb{C}^{m,p}(\Omega)$.

Poichè lo spazio vettoriale sostegno di $\mathbb{C}^{m,p}(\Omega)$ è $C^m(\bar{\Omega})$, solitamente si dice che $H^{m,p}(\Omega)$ è il completamento di $C^m(\bar{\Omega})$ con la norma di $W^{m,p}(\Omega)$.

Notiamo che il completamento $H^{m,p}(\Omega)$ non è lo spazio normato completo $\hat{S}$ costruito nella dimostrazione del teorema (4.1). Un elemento X di $\hat{S}$ è una classe di successioni; un rappresentante di X è una successione di Cauchy di $\mathbb{C}^{m,p}(\Omega)$, indichiamola $(x_n)_{n\in\mathbb{N}}$, e risulta:

$$\|X\|_{\hat{S}} = \lim \|x_n\|_{\mathbb{C}^{m,p}(\Omega)} .$$

Osserviamo che, se A è l'isomorfismo isometrico sopra introdotto, la successione $(Ax_n)_{n\in\mathbb{N}}$ è convergente in $W^{m,p}(\Omega)$, poichè:

$$(x_n)_{n\in\mathbb{N}} \text{ succ. di Cauchy in } \mathbb{C}^{m,p}(\Omega) \Longleftrightarrow (Ax_n)_{n\in\mathbb{N}} \text{ succ. di Cauchy in } W^{m,p}(\Omega)$$

e lo spazio $W^{m,p}(\Omega)$ è completo. Detto allora L l'operatore che ad ogni $X \in \hat{S}$ associa il limite di Ax_n in $W^{m,p}(\Omega)$, tenendo presente quanto detto nella nota (1) di pag. 163 si ha:

$$\|LX\|_{H^{m,p}(\Omega)} = \|LX\|_{W^{m,p}(\Omega)} = \|\lim Ax_n\|_{W^{m,p}(\Omega)} = \lim \|Ax_n\|_{W^{m,p}(\Omega)} = \lim \|x_n\|_{\mathbb{C}^{m,p}(\Omega)} = \|X\|_{\hat{S}} .$$

Abbiamo così constatato che esiste un isomorfismo isometrico di $\hat{S}$ su $H^{m,p}(\Omega)$, appunto l'operatore L, in conformità con quanto afferma il teorema (4.1).

---

(1) Quando è stato introdotto lo spazio $\mathbb{C}^{m,p}(\Omega)$, il sottospazio H è stato identificato, come spazio vettoriale, con lo spazio vettoriale $C^m(\bar{\Omega})$: ciò allo scopo di semplificare il discorso.

5. Operatori compatti.

Siano S, S_1 spazi vettoriali normati. Se S_1 è a dimensione finita, ogni operatore lineare e continuo di S in S_1 trasforma insiemi limitati in insiemi precompatti: basta infatti osservare che, in virtù del teorema (2.5) di pag. 141, in uno spazio a dimensione finita gli insiemi precompatti sono gli insiemi limitati, e ricordare che un operatore lineare e continuo trasforma insiemi limitati in insiemi limitati [cfr. (2.1) pag. 138].

Se S_1 non è a dimensione finita, esistono operatori lineari e continui di S in S_1 che non trasformano insiemi limitati in insiemi precompatti; basta infatti considerare l'operatore identico di S_1 in S_1, che trasforma l'insieme limitato:

$$\Gamma = \{ x \in S_1 : \|x\| = 1 \}$$

in un insieme (lo stesso Γ) che non è precompatto (*).

La proprietà di trasformare insiemi limitati in insiemi precompatti gioca un ruolo importante nelle applicazioni: ciò induce a prendere in considerazione quei particolari operatori lineari e continui che la verificano, e quindi ad introdurre la seguente definizione.

Con riferimento agli spazi vettoriali normati S, S_1, sia

$$A : S \to S_1$$

un operatore lineare. Si dice che A è un <u>operatore compatto</u>, quando esso trasforma insiemi limitati in insiemi precompatti; ossia quando, qualunque sia $X \subseteq S$, sussiste l'implicazione:

(1) $$X \text{ limitato} \Longrightarrow A(X) \text{ precompatto}.$$

In tale definizione si fa riferimento ad un operatore lineare, senza supporre che esso sia continuo; ma <u>se un operatore lineare A è compatto, esso è anche continuo</u>: basta infatti osservare che:

$$A(X) \text{ precompatto} \Longleftrightarrow \overline{A(X)} \text{ compatto} \Longrightarrow \overline{A(X)} \text{ limitato} \quad (2)$$

e che pertanto, essendo $A(X) \subseteq \overline{A(X)}$, in virtù della (1) A trasforma insiemi limitati in insiemi limitati, ossia A è continuo [cfr. (2.1) pag. 138].

(*) Si dimostra che: <u>detto S uno spazio vettoriale normato, l'insieme $\Gamma = \{x \in S : \|x\| = 1\}$ è compatto</u> (ossia precompatto, in quanto Γ è chiuso) <u>se e solo se S è a dimensione finita</u>.

(2) Cfr. (1.6) pag. 80.

Quindi la proprietà che A sia un operatore compatto significa, per così dire, "qualcosa in più" della continuità : è questa la ragione per cui un operatore lineare compatto si dice anche completamente continuo.

La nozione introdotta si può riferire anche ad una applicazione di uno spazio metrico in un altro, ma in tal caso la continuità va presupposta e si preferisce la locuzione "completa continuità". Precisamente, un'applicazione f di uno spazio metrico in un altro si dice completamente continua quando è continua ed inoltre trasforma insiemi limitati in insiemi precompatti.

Utile è l'osservazione seguente :

(5.1) Siano A, B operatori lineari tra spazi vettoriali normati. Se A è compatto e B è continuo, allora l'operatore AB (purché abbia senso) è compatto; se inoltre D_0 è chiuso, anche BA (purché abbia senso) è compatto.

Dim. Se $X \subseteq D_{AB}$ è limitato, $B(X)$ è limitato, e quindi $AB(X)$ è precompatto. Pertanto AB è un operatore compatto.

Se $X \subseteq D_{BA}$ è limitato, $A(X)$ è precompatto, ossia $\overline{A(X)}$ è compatto; se D_B è chiuso, $\overline{A(X)}$ è certo contenuto in D_B, e $B(\overline{A(X)})$ è compatto perchè B è continuo (1) . Essendo $BA(X)$ contenuto in un compatto, $BA(X)$ è precompatto. Pertanto BA è un operatore compatto.

Una importante proprietà degli operatori compatti è espressa dal teorema :

(5.2) Siano: S uno spazio vettoriale normato, $A: S \to S$ un operatore lineare e I l'identità in S .

Se A è compatto, qualunque sia $\lambda \in K-\{0\}$ l'operatore :

$$A_\lambda = A - \lambda I \tag{2}$$

ha il codominio chiuso, e il nucleo a dimensione finita.

Dim. Ovviamente basta dimostrare l'asserto per l'operatore $A_1 = A - I$.

Procedendo per assurdo, ammettiamo che $A_1(S)$ non sia chiuso. In virtù della (4.8) di pag.109 , esiste allora una successione convergente di punti di $A_1(S)$, diciamola $(y_n)_{n\in N}$, il cui limite y_0 non appartiene ad $A_1(S)$.

Ovviamente $y_0 \neq 0$ [giacchè $0 \in A_1(S)$] . Supposto che sia:

$$y_n = A_1 x_n \quad \forall n \in N, \tag{3}$$

esiste $\nu \in N$ tale che $x_n \notin \operatorname{Ker} A_1 \ \forall n > \nu$: se così non fosse, dalla successione $(x_n)_{n\in N}$ si potrebbe estrarre una successione $(x_{k_n})_{n\in N}$ di punti di $\operatorname{Ker} A_1$,

(1) Cfr. (7.5) pag.22 .

e per la (3) si avrebbe $y_{k_n}=0 \quad \forall n\in N$, il che é assurdo in quanto la successione $(y_{k_n})_{n\in N}$ ha per limite $y_0\neq 0$. Possiamo quindi supporre, senza ledere la generalità, che $x_n \notin \ker A_1$, per ogni $n\in N$: essendo $\ker A_1$ chiuso [1] , per ogni $n\in N$ la distanza d_n di x_n da $\ker A_1$ é positiva.

Per ogni $n\in N$, esiste allora almeno un punto x'_n tale che:

$$(4) \qquad x'_n \in \ker A_1 \quad , \quad \|x_n - x'_n\| < 2d_n .$$

Posto:

$$(5) \qquad t_n = \|x_n - x'_n\| \quad \forall n\in N ,$$

vogliamo provare che:

$$(6) \qquad \lim_{n\to\infty} t_n = +\infty ,$$

e a tale fine ammettiamo che la (6) sia falsa; allora dalla successione $(t_n)_{n\in N}$ si può estrarre una successione limitata , diciamola $(t_{k_n})_{n\in N}$. Ricordando la (5), la successione $(x_{k_n} - x'_{k_n})_{n\in N}$ é dunque limitata : essendo A compatto , ne segue che l'insieme degli elementi della successione $(A(x_{k_n} - x'_{k_n}))_{n\in N}$ é precompatto, e che pertanto da tale successione si può estrarre una successione convergente [2] : possiamo supporre, senza ledere la generalità, che la successione considerata sia essa stessa convergente. Essendo:

$$[x_{k_n} - x'_{k_n}] - A(x_{k_n} - x'_{k_n}) = -y_{k_n} \quad \forall n\in N \quad [3],$$

e quindi:

$$x_{k_n} - x'_{k_n} = A(x_{k_n} - x'_{k_n}) - y_{k_n} ,$$

anche la successione $(x_{k_n} - x'_{k_n})_{n\in N}$ é convergente e, detto w il suo limite, si ha:

$$w = Aw - y_0$$

ossia $y_0 = Aw - w = A_1 w$, il che é assurdo in quanto $y_0 \notin A_1(S)$.

Pertanto la (6) é acquisita.

Poiché $x'_n \in \ker A_1$ e $x_n \notin \ker A_1$, si ha $t_n \neq 0 \ \forall n\in N$.

Posto:

$$(7) \qquad z_n = \frac{x_n - x'_n}{t_n} ,$$

si ha:

$$z_n - Az_n = \frac{1}{t_n}\left[(x_n - x'_n) - A(x_n - x'_n)\right] = -\frac{1}{t_n} y_n$$

e quindi:

$$(8) \qquad z_n = Az_n - \frac{1}{t_n} y_n .$$

Poiché $(z_n)_{n\in N}$ é limitata [4] , l'insieme degli elementi di $(Az_n)_{n\in N}$ é precompatto, e quindi da tale successione si può estrarre una successione convergente: non si lede la generalità supponendo che la stessa successione $(Az_n)_{n\in N}$ sia convergen-

---

(1) Tenendo presente che A é continuo (in quanto compatto), in base alla (1.8) di pag. 109 si verifica immediatamente che il nucleo di A_1 é chiuso.

(2) Cfr. (2.1) pag.111.

(3) Per la prima delle (4) risulta $Ax'_{k_n} - x'_{k_n} = 0$, e per la (3) $Ax_{k_n} - x_{k_n} = y_{k_n}$.

(4) Da (5) e (7) segue che $\|z_n\| = 1 \ \forall n\in N$.

te . Dalla (8), in base alla (6) e al fatto che $(y_n)_{n\in N}$ è convergente, si deduce:

$$\lim_{n\to\infty} z_n = \lim_{n\to\infty} A z_n = A\left(\lim_{n\to\infty} z_n\right) ,$$

e quindi, detto z_0 il limite al primo membro, si ha:

(9) $$z_0 = A z_0 \quad , \text{ ossia } \quad z_0 \in \ker A_1 .$$

Osserviamo ora che per ogni $n \in N$, si ha, tenendo presente la (7), la (5) e la seconda delle (4):

(10) $$\| x_n - x'_n - t_n z_0 \| = \| t_n z_n - t_n z_0 \| = t_n \| z_n - z_0 \| < 2 d_n \| z_n - z_0 \| ;$$

poiché, per la (9) e per la prima delle (4), il punto $x'_n + t_n z_0$ appartiene a $\ker A_1$, per la definizione di d_n il primo membro della (10) è $\geq d_n$, e quindi la (10) stessa implica:

$$d_n < 2 d_n \| z_n - z_0 \| \qquad \forall n \in N$$

ossia:

$$1 < 2 \| z_n - z_0 \| \qquad \forall n \in N .$$

Ma ciò è assurdo, in quanto z_0 è il limite della successione $(z_n)_{n\in N}$.

Quanto sopra dimostra che $A_1(S)$ è chiuso.

Proviamo ora che $\ker A_1$ (sottospazio di S) è a dimensione finita.

A tale scopo, consideriamo il sottoinsieme di $\ker A_1$:

$$\Gamma = \{ x \in \ker A_1 : \|x\| = 1 \} .$$

Qualunque sia la successione $(x_n)_{n\in N}$ di punti di Γ , essendo $(x_n)_{n\in N}$ limitata e A compatto, dalla successione $(Ax_n)_{n\in N}$ si può estrarre una successione convergente. Ma $x_n \in \ker A_1 \ \forall n \in N$, ossia:

$$A x_n = x_n \qquad \forall n \in N$$

e quindi anche dalla successione $(x_n)_{n\in N}$ si può estrarre una successione convergente.

Abbiamo con ciò dimostrato che il sottoinsieme Γ dello spazio vettoriale normato $\ker A_1$ è sequenzialmente precompatto, ossia precompatto [(2.1) pag. 111] . In virtù del teorema enunciato nella nota (1) di pag. 165 , ne consegue che lo spazio $S_1 = \ker A_1$ è a dimensione finita.

Esempi.

1) Equazioni integrali di Fredholm.

Sia $K(s,t)$ una funzione complessa di due variabili, elemento dello spazio $L^2(\Omega_1 \times \Omega_2)$, con Ω_1, Ω_2 aperti limitati di R^k. Considerando $K(s,t)$ come funzione di una sola delle variabili, per quasi tutti i valori dell'altra variabile essa appartiene ad L^2, e si ha:

$$(11)\qquad \iint_{\Omega_1\times\Omega_2} |K(\sigma,t)|^2 d\sigma\,dt = \int_{\Omega_1} d\sigma \int_{\Omega_2} |K(\sigma,t)|^2 dt = \int_{\Omega_2} dt \int_{\Omega_1} |K(\sigma,t)|^2 d\sigma \quad (1).$$

Ne segue che qualunque sia $x \in L^2(\Omega_2)$ [$x \in L^2(\Omega_1)$], la funzione:

$$(12)\qquad y_1(\sigma) = \int_{\Omega_2} K(\sigma,t)\,x(t)\,dt \qquad \left[\; y_2(t) = \int_{\Omega_1} K(\sigma,t)\,x(\sigma)\,d\sigma \;\right]$$

è definita q.o. in Ω_1 [in Ω_2], ed appartiene ad $L^2(\Omega_1)$ [ad $L^2(\Omega_2)$]; infatti, con riferimento ad es. ad y_1 si ha, per la disuguaglianza di Schwarz [cfr. (3) pag. 54 con $p=p'=2$]

$$|y_1(\sigma)|^2 = \left(\left| \int_{\Omega_2} K(\sigma,t)\,x(t)\,dt \right| \right)^2 \le \int_{\Omega_2} |K(\sigma,t)|^2 dt \cdot \int_{\Omega_2} |x(t)|^2 dt .$$

e quindi, per la (11) (2):

$$(13)\qquad \int_{\Omega_1} |y_1(\sigma)|^2 d\sigma \le \left(\int_{\Omega_1} d\sigma \int_{\Omega_2} |K(\sigma,t)|^2 dt \right) \left(\int_{\Omega_2} |x(t)|^2 dt \right) = \|K\|^2_{L^2(\Omega_1\times\Omega_2)} \cdot \|x\|^2_{L^2(\Omega_2)}$$

Consideriamo dunque l'applicazione

$$T_1 : L^2(\Omega_2) \longrightarrow L^2(\Omega_1) \qquad \left[\; T_2 : L^2(\Omega_1) \to L^2(\Omega_2) \;\right]$$

che ad ogni $x \in L^2(\Omega_2)$ [$x \in L^2(\Omega_1)$] associa la funzione y_1 [y_2] definita dalla (12). Si tratta ovviamente di un operatore lineare, che è anche continuo in quanto risulta per la (13) [per l'analoga della (13) relativa ad y_2]:

$$\|T_1 x\| \le \|K\| \cdot \|x\| \qquad \left[\; \|T_2 x\| \le \|K\| \cdot \|x\| \;\right].$$

Gli operatori T_1 e T_2 si chiamano <u>operatori integrali, generati dal "nucleo" K</u>: per distinguerli è opportuno dire che T_1 è generato dal nucleo $K(\sigma,t)$ <u>con variabile di integrazione t</u>, e T_2 dal nucleo $K(\sigma,t)$ <u>con variabile di integrazione σ</u> (facendo così riferimento alla (12), dove $y_1 = T_1 x$ [$y_2 = T_2 x$]).

Supponiamo ora $\Omega_1 = \Omega_2 = \Omega$. Per ogni fissato $y \in L^2(\Omega)$ l'equazione nell'incognita $x \in L^2(\Omega)$:

$$(14)\qquad T_1 x - \lambda x = y \qquad \left[\; T_2 x - \lambda x = y \;\right]$$

dove λ è un fissato numero complesso, si chiama un'<u>equazione integrale di Fredholm, di nucleo K</u>; se il "termine noto" y è nullo, l'equazione si dice <u>omogenea</u>. Naturalmente l'equazione omogenea "associata alla (14)" è quella che da essa si ottiene ponendovi $y=0$.

Orbene, vedremo in seguito che <u>T_1 e T_2 sono operatori compatti</u> [pag. 201].

In virtù del teorema (5.2) di pag. 166, l'equazione omogenea associata alla (14), per ogni fissato $\lambda \neq 0$, o ammette soltanto la soluzione nulla, oppure ammette infinite

---

(1) Tale asserzione costituisce il <u>teorema di Fubini</u>, che però si riferisce ad una funzione $F(\sigma,t) \in L^1(\Omega_1\times\Omega_2)$ [nel nostro caso, $F(\sigma,t) = |K(\sigma,t)|^2$]; il teorema assicura **che l'integrale** di F rispetto ad una delle variabili è sommabile rispetto all'altra variabile (questa circostanza è stata da noi sottintesa nell'uguaglianza (11)), **e che vale l'uguaglianza**.

Il teorema citato generalizza la formula di riduzione degli integrali doppi.

(2) Si tenga presente quanto detto nella nota (3) a pag. 54.

soluzioni che costituiscono un sottospazio di $L^2(\Omega)$ a dimensione finita; perciò nella seconda eventualità, detta $x_1, \dots, x_n$ una base del sottospazio suo menzionato, tutte le soluzioni della omogenea associata si ottengono dalla formula $x = c_1 x_1 + \dots + c_n x_n$ al variare delle costanti $c_1, \dots, c_n$.

Il citato teorema (5.2) assicura poi che per ogni $\lambda \neq 0$ l'operatore $J_1 - \lambda I$ $[J_2 - \lambda I]$ ha il codominio chiuso: da tale proprietà, come vedremo nel seguito, si traggono delle conclusioni riguardanti l'equazione completa.

2) L'immersione di $C^{0,\lambda}(T)$ in $C^0(T)$:

$$J_0 : x \in C^{0,\lambda}(T) \to x \in C^0(T)$$

è un operatore compatto [cfr. es. 6 pag. 101].

L'immersione di $C^{m,\lambda}(T)$ in $C^m(T)$:

$$J_m : x \in C^{m,\lambda}(T) \to x \in C^m(T)$$

è un operatore compatto [cfr. es. 7 pag. 101].

Se T è uno pseudodominio convesso di R^n, anche l'immersione di $C^{m+1}(T)$ in $C^m(T)$:

$$J : x \in C^{m+1}(T) \to x \in C^m(T)$$

è un operatore compatto. Infatti, detta y una qualunque delle derivate m^{me} di $x \in C^{m+1}(T)$, per il teorema di Lagrange si ha, qualunque siano $t', t'' \in T$:

$$|y(t') - y(t'')| \leq \sup |\operatorname{grad} y| \cdot |t' - t''| \leq \|y\|_{C^1(T)} \cdot |t' - t''| ,$$

sicchè:

$$y \in C^0(T) \quad , \quad [y]_1 \leq \|y\|_{C^1(T)} .$$

Pertanto si ha $C^{m+1}(T) \subset C^{m,1}(T)$, e $\|x\|_{C^{m,1}(A)} \leq 2 \|x\|_{C^{m+1}(A)} \quad \forall x \in C^{m+1}(T)$. Ne segue che un insieme limitato in $C^{m+1}(T)$ è limitato anche in $C^{m,1}(T)$ e quindi è precompatto in $C^m(T)$.

3) Se S è uno spazio vettoriale normato di dimensione infinita, l'identità in S non è un operatore compatto.

Basta ricordare il teorema enunciato nella nota (1) a pag. 165.

4) Se S è uno spazio vettoriale normato di dimensione infinita, un operatore compatto di S su S non può essere dotato di inverso continuo.

Infatti, se per assurdo l'operatore compatto A fosse dotato di inverso continuo, per il teorema (5.1) di pag. 166 l'operatore $A \cdot A^{-1} = I$ sarebbe compatto, il che è assurdo per il precedente esempio.

6. Duale topologico di uno spazio vettoriale topologico. Duale forte di uno spazio vettoriale normato.

D'ora in avanti il termine "funzionale" sarà riferito sia ad un funzionale reale, sia ad un funzionale complesso; per un funzionale definito in uno spazio vettoriale Σ, se Σ è sul campo reale [complesso] si intenderà che si tratta di un funzionale reale [complesso].

Ciò premesso, ricordiamo che per ogni spazio vettoriale Σ abbiamo denotato con Σ^+ il duale algebrico di Σ [cfr. pag. 69]: se Σ è sul campo reale, Σ^+ è lo spazio vettoriale costituito dai funzionali lineari definiti in Σ; se Σ è sul campo complesso, Σ^+ è lo spazio vettoriale costituito dai coniugati dei funzionali lineari definiti in Σ, o ciò che è lo stesso dai funzionali definiti in Σ i cui coniugati sono funzionali lineari. Naturalmente in entrambi i casi si può dire che gli elementi di Σ^+ sono i coniugati dei funzionali lineari definiti in Σ.

Consideriamo ora uno spazio vettoriale topologico $S=(\Sigma,\tau)$.

Il sottoinsieme di Σ^+ costituito dagli elementi di Σ^+ che sono continui in S è ovviamente un sottospazio vettoriale di Σ^+, e si chiama il duale topologico di S. In tale locuzione si fa riferimento ad S e non solo al suo sostegno Σ, in quanto il duale topologico dipende anche dalla topologia τ associata a Σ: se si considera lo spazio vettoriale topologico $S_1=(\Sigma,\tau_1)$ col sostegno uguale a quello di S ma con la topologia τ_1 diversa da τ, il duale topologico di S_1 è diverso dal duale topologico di S.

Il duale topologico di S sarà da noi denotato con S^* (1).

L'operazione che ad ogni coppia $(u,x)\in S^*\times S$ associa il numero (reale se S è reale, complesso se S è complesso):

$$(1) \qquad \langle u, x\rangle ,$$

ossia il valore del funzionale u nel punto x, si chiama la dualità tra S ed S^*: si tratta di una forma sesquilineare (bilineare se S è reale), che è la

(1) Avvertiamo che spesso il duale topologico di S viene denotato con S'.

restrizione ad $S^* \times S$ della "dualità tra Σ e Σ^*" introdotta a pag. 69 (1).

Per ogni fissato $x \in S$, il funzionale definito in S^*:

$$u \in S^* \longrightarrow \langle u, x \rangle \tag{2}$$

è ovviamente lineare, e si denota col simbolo:

$$\langle \cdot, x \rangle \tag{2}$$

Naturalmente, l'analogo simbolo per ogni fissato $u \in S^*$:

$$\langle u, \cdot \rangle \tag{2'}$$

rappresenta il funzionale u, che è continuo e gode delle proprietà:

$$\langle u, x_1 + x_2 \rangle = \langle u, x_1 \rangle + \langle u, x_2 \rangle \quad , \quad \langle u, \lambda x \rangle = \overline{\lambda} \langle u, x \rangle \, .$$

Infine il simbolo $\langle \cdot, \cdot \rangle$ denota il funzionale che abbiamo chiamato la dualità tra S ed S^*.

Quando S è complesso, la u che figura nel simbolo (1) denota non un funzionale lineare, ma un funzionale il cui coniugato è lineare; avvertiamo che talvolta, per evitare inutili complicazioni formali, adopereremo il simbolo (1) anche se u è un funzionale lineare. Pur essendo banale, rileviamo che risulta $|\langle u, x \rangle| = |\langle \bar{u}, x \rangle| \quad \forall u \in S^*$ e $\forall x \in S$.

Osservazione.

Alcuni Autori definiscono il duale topologico di S come lo spazio vettoriale costituito dai funzionali lineari definiti in Σ e continui in S, anziché dai coniugati di detti funzionali. Con tale definizione, nel caso S complesso la dualità tra S ed S^* è una forma bilineare e non una forma sesquilineare.

A noi sembra invece opportuno che la dualità $\langle \cdot, \cdot \rangle$ sia dello stesso tipo del prodotto scalare $(\cdot, \cdot)$ in uno spazio pre-hilbertiano (il motivo sarà chiaro nel seguito), e il prodotto scalare è una forma sesquilineare.

Ci si può domandare perché il prodotto scalare è stato definito come forma sesquilineare e non come forma bilineare. Si pensi al prodotto scalare in C^n [pag. 98]: se si ponesse $(x,y) = \sum_{i=1}^{n} x_i y_i$ anche in C^n, per $x = y$ si avrebbe $\sum_{i=1}^{n} x_i^2$ e non si otterrebbe il quadrato della norma, che è invece $\sum_{i=1}^{n} |x_i|^2$.

---

(1) Gli elementi di S^* sono particolari elementi di Σ^+, e gli elementi di S sono gli elementi di Σ (giacché questi ultimi si sogliono chiamare punti di S, o elementi di S): perciò $S^* \times S$ è un sottospazio di $\Sigma^+ \times \Sigma$.

Se S è uno spazio vettoriale normato a dimensione finita (necessariamente completo, cfr. nota (1) a pag. 141), in virtù di quanto stabilito nell'es. 3 di pag. 145 risulta $\Sigma' \subseteq S^*$, e quindi si ha l'uguaglianza in quanto $S^* \subseteq \Sigma'$. Si dimostra che viceversa, se S è uno spazio vettoriale normato e risulta $S^* = \Sigma'$, lo spazio è a dimensione finita. Vale dunque il teorema:

(6.1) Se $S = (\Sigma, \nu)$ è uno spazio vettoriale normato, il suo duale topologico S^* coincide col duale algebrico di Σ quando e solo quando Σ è a dimensione finita.

Il duale topologico S^* di uno spazio vettoriale topologico S non è munito di alcuna topologia, per cui esso è semplicemente uno spazio vettoriale. Peraltro ad S^* si sogliono associare diverse topologie, alcune delle quali saranno introdotte nel seguito con riferimento al caso in cui S è uno spazio vettoriale normato. In tal caso ad S^* si suole associare anche una norma, che ci proponiamo ora di definire.

Supponiamo dunque che S sia uno spazio vettoriale normato.

Qualunque sia $u \in S^*$, poichè u è un funzionale lineare e continuo in S esiste, per la (2.1) di pag. 138, un numero positivo ℓ tale che:

$$|u(x)| \leq \ell \|x\| \qquad \forall x \in S \quad (1)$$

e quindi tale che:

$$\frac{|u(x)|}{\|x\|} \leq \ell \qquad \forall x \in S - \{0\}.$$

Pertanto possiamo considerare il numero reale non negativo:

$$(4) \qquad \nu^*(u) = \sup_{x \in S - \{0\}} \frac{|u(x)|}{\|x\|}.$$

La (4) definisce un'applicazione ν^* di S^* in $[0, +\infty[$ che, come facilmente si verifica, è una norma per S^*.

Lo spazio vettoriale normato (S^*, ν^*), che si denota semplicemente con S^*, si chiama il duale forte di S: dunque il duale forte di S è lo spazio vettoriale normato avente per sostegno il duale topologico di S e per norma quella definita mediante la (4). Avvertiamo che l'attributo "forte" spesso si sottintende, purchè non vi sia possibilità di equivoci.

---

(1) Abbiamo qui invocato la (2.1) con $S_1 = K$, e quindi la norma di $\bar{u}(x) \in S_1$ è il modulo di $\bar{u}(x)$, che è uguale a quello di $u(x)$.

Dopo quanto si è detto, la (4) si può scrivere:

$$(4') \qquad \|u\|_{S^*} = \sup_{x \in S-\{0\}} \frac{|u(x)|}{\|x\|} \qquad \forall u \in S^*$$

e risulta:

$$(5) \qquad |u(x)| \leq \|u\|_{S^*} \cdot \|x\| \qquad \forall u \in S^* \text{ e } \forall x \in S.$$

Adoperando il simbolo (1) si ha dunque:

$$(5') \qquad |\langle u, x \rangle| \leq \|u\| \cdot \|x\| \qquad \forall u \in S^* \text{ e } \forall x \in S$$

avendo denotato con $\|u\|$, naturalmente, la norma di u in S^*.

Si osservi che il secondo membro della (4') rimane inalterato se ivi si sostituisce u con $\bar{u}$: perciò quando S è complesso, se u è un funzionale lineare e continuo in S (poichè non u, ma $\bar{u}$ è un elemento di S^*), talvolta scriveremo $\|u\|_S$, in luogo di $\|\bar{u}\|_{S^*}$.

Si noti l'analogia tra la (5') e la disuguaglianza di Schwarz, che riguarda il prodotto scalare in uno spazio pre-hilbertiano. La dualità $\langle \cdot, \cdot \rangle$ tra S ed S^*, e il prodotto scalare $(\cdot, \cdot)$ in uno spazio pre-hilbertiano, sono entrambi delle forme sesquilineari (bilineari se lo spazio è reale), la prima in

$$(6) \qquad S^* \times S$$

e la seconda in

$$(6') \qquad S \times S$$

Orbene, vedremo nel numero seguente che uno spazio di Hilbert è equivalente al suo duale; inoltre se S è uno spazio di Hilbert, identificando S con S^* non solo si vengono ad identificare gli spazi (6) e (6'), ma la dualità, come vedremo, si riduce proprio al prodotto scalare. Perciò la dualità $\langle \cdot, \cdot \rangle$ tra uno spazio vettoriale normato S e il suo duale S^* si presenta come una operazione che generalizza il prodotto scalare.

Le considerazioni svolte giustificano le seguenti definizioni. Se S è uno spazio vettoriale normato, qualunque siano $x \in S$ ed $u \in S^*$, se risulta:

$$\langle u, x \rangle = 0$$

si dice che <u>u ed x sono ortogonali</u>. Di conseguenza si dice che <u>l'elemento u di S^* è ortogonale al sottoinsieme X di S</u> quando $\langle u, x \rangle = 0 \ \forall x \in X$, e che <u>l'elemento x di S è ortogonale al sottoinsieme U di S^*</u> quando $\langle u, x \rangle = 0 \ \forall u \in U$.

Dimostriamo ora due teoremi concernenti lo spazio duale.

(6.2) <u>Qualunque sia lo spazio vettoriale normato S, il duale di S è uno spazio di Banach.</u>

Dim. Per semplicità supporremo S reale, ma il ragionamento si può

ripetere, salvo modifiche puramente formali, anche nel caso S complesso.

Sia $(u_n)_{n\in N}$ una successione di Cauchy di S^*. Per ogni $\varepsilon>0$, esiste $m\in N$ tale che:

$$\|u_p-u_q\|<\varepsilon \quad \forall p\geq m \text{ e } \forall q\geq m .$$

Qualunque sia $x\in S-\{0\}$ si ha allora [cfr.(5)] :

(7) $$|u_p(x)-u_q(x)|<\varepsilon\|x\| \qquad \forall p\geq m \text{ e } \forall q\geq m$$

e quindi la successione di numeri reali $(u_n(x))_{n\in N}$ soddisfa alla condizione di Cauchy: pertanto, tale successione è convergente.

Evidentemente il funzionale:

$$u : x\in S \longrightarrow \lim_{n\to\infty} u_n(x)$$

è lineare; d'altro canto dalla (7), passando al limite per $p\to\infty$, si deduce:

(8) $$|u(x)-u_q(x)|\leq\varepsilon\|x\| \qquad \forall q\geq m \text{ e } \forall x\in S$$

e quindi risulta:

$$|u(x)|\leq|u(x)-u_m(x)|+|u_m(x)|\leq(\varepsilon+\|u_m\|)\cdot\|x\| .$$

Dunque u è anche continuo, e pertanto $u\in S^*$. Poichè la (8) implica:

$$\|u-u_m\|\leq\varepsilon \qquad \forall n\geq m$$

si conclude che la successione $(u_n)_{n\in N}$ è convergente (in S^*) ad u. Poichè ogni successione di Cauchy di S^* è convergente, S^* è completo.

- - - - - - - - - - - - - -

(6.3) <u>Ogni spazio vettoriale normato a dimensione finita è isomorfo al suo duale.</u>

- - - - - - - - - - - - - -

Dim. Sia S uno spazio vettoriale normato di dimensione finita m, e sia Σ il suo sostegno.

Per la (6.1) di pag.173 lo spazio S^* coincide con Σ^+, e questo è isomorfo algebricamente a K^m [es.1 pag.70] : dunque anche S^* ha dimensione finita m.

Per la (2.5) di pag.141 gli spazi S, S^* sono entrambi isomorfi a K^m, e pertanto sono isomorfi.

- - - - - - - - - - - - - -

Supposto che lo spazio vettoriale normato S sia sul campo reale, vogliamo infine stabilire un teorema che fornisce una interpretazione geometrica della continuità di un funzionale lineare definito in S, nonchè della norma di u in S^*. A tale scopo ricordiamo [pag.63] che se u è un funzionale lineare definito in S, non identicamente nullo, cioè se u è un elemento non nullo del duale algebrico, per ogni $\alpha\in R$ il sottoinsieme di S:

(9) $$\{x\in S : u(x)=\alpha\}$$

è un iperpiano di S (<u>iperpiano di equazione $u(x)=\alpha$</u>). Sussiste il teorema:

(6.4) Sia u un funzionale lineare definito in S, non identicamente nullo. Fissato $\alpha \in R$, il funzionale u è continuo se e solo se l'iperpiano Π di S di equazione $u(x)=\alpha$ è un insieme chiuso. Se u è continuo, per ogni punto $x_0 \notin \Pi$ il rapporto:

$$\frac{|u(x_0)-\alpha|}{\|u\|_{S^*}} \tag{10}$$

è uguale alla distanza di x_0 da Π; di conseguenza risulta:

$$\|u\|_{S^*} = \frac{1}{\delta(u)} ,$$

dove $\delta(u)$ è la distanza del punto 0 di S dall'iperpiano di equazione $u(x)=1$.

Dim. Se u è continuo, per la (7.4) di pag. 21 l'insieme (9), uguale ad $\bar{u}(\{\alpha\})$, è chiuso.

Viceversa supponiamo che l'iperpiano Π sia chiuso, e consideriamo un punto $x_0 \notin \Pi$, sicché risulta $u(x_0) \neq \alpha$. Per fissare le idee, sia $u(x_0) < \alpha$.

Detto r un numero positivo minore della distanza $d(x_0,\Pi)$ di x_0 da Π, il cerchio chiuso Γ di centro x_0 e raggio r non ha punti in comune con Π. Mostriamo che risulta:

$$u(y) < \alpha \quad \forall y \in \Gamma , \tag{11}$$

e a tale scopo ammettiamo che esista $y_0 \in \Gamma$ tale che $u(y_0) > \alpha$. Allora esiste un punto di $\Gamma \cap \Pi$ (assurdo), giacché ha i punti del segmento di estremi x_0, y_0, contenuto in Γ, ce n'è uno nel quale il funzionale u assume il valore α: infatti risolvendo l'equazione $u((1-t)x_0 + t y_0) = \alpha$ si trova $t = [\alpha - u(x_0)] / [u(y_0) - u(x_0)]$, e questo numero è minore di 1 perché $\alpha < u(y_0)$, ed è positivo perché $u(x_0) < \alpha$ e $u(y_0) > \alpha$.

Orbene, qualunque sia $x \in S-\{0\}$ il punto $y = x_0 \pm r \frac{x}{\|x\|}$ appartiene a Γ, e quindi per la (11) si ha:

$$u\left(x_0 \pm r \frac{x}{\|x\|}\right) < \alpha$$

ossia $u(x_0) \pm \frac{r}{\|x\|} u(x) < \alpha$, da cui:

$$\frac{|u(x)|}{\|x\|} < \frac{\alpha - u(x_0)}{r} \quad \forall x \in S-\{0\} . \tag{12}$$

Poiché per la (12) $\exists l > 0$ tale che $|u(x)| \leq l \|x\| \ \ \forall x \in S$ (la disuguaglianza è banale se $x=0$), per la (2.1) di pag. 138 il funzionale lineare u è continuo.

Dalla (12) segue poi, per la definizione della norma in S^*:

$$\|u\|_{S^*} \leq \frac{\alpha - u(x_0)}{r} ,$$

ossia:

$$r \leq \frac{\alpha - u(x_0)}{\|u\|_{S^*}}$$

e quindi, essendo r un qualunque numero positivo minore di $d(x_0,\Pi)$, risulta:

$$d(x_0,\Pi) \leq \frac{\alpha - u(x_0)}{\|u\|_{S^*}} . \tag{13}$$

Osserviamo ora che, detto δ il rapporto (10), per $\|x-x_0\| < \delta$ si ha, tenendo presente la (5') di pag. 174:

$$u(x) = u(x-x_0) + u(x_0) \leq \|u\| \cdot \|x-x_0\| + u(x_0) < \|u\| \frac{\alpha - u(x_0)}{\|u\|} + u(x_0) = \alpha , \quad \text{quindi } x \notin \Pi .$$

Ne segue che per ogni $x \in \Pi$ risulta $\|x-x_0\| \geq \delta$, e perciò $d(x_0,\Pi) \geq \delta$. In virtù della (13) si ha dunque $d(x_0,\Pi) = \delta$ e l'assunto è dimostrato.

È possibile dimostrare, ma su ciò non ci soffermiamo, che se il funzionale lineare u definito in S non è continuo, qualunque sia $\alpha \in R$ l'iperpiano di equazione $u(x)=\alpha$ è denso in S.

Esempi.

1. Consideriamo il funzionale lineare:

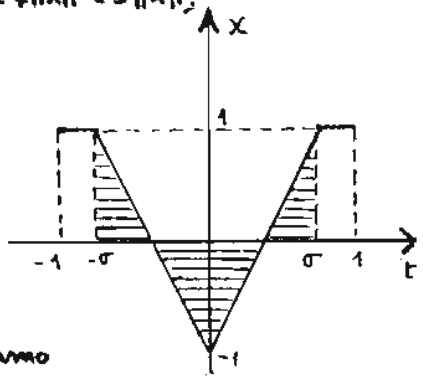
$$u : x \in C^0([-1,1]) \to \int_{-1}^{1} x(t)\,dt - x(0).$$

Qualunque sia $x \in C^0([-1,1])$ si ha:

$$|u(x)| \le \int_{-1}^{1} |x(t)|\,dt + |x(0)| \le \int_{-1}^{1} \|x\|\,dt + \|x\| = 3\|x\|,$$

e quindi u è continuo, e risulta:

$$(1) \qquad \frac{|u(x)|}{\|x\|} \le 3 \qquad \forall x \in C^0([-1,1]) - \{0\}.$$



Mostriamo che 3 è l'estremo superiore del primo membro della (1). A tale scopo, fissato un numero positivo $\sigma < 1$, consideriamo la funzione x_σ definita in $[-1,1]$ avente il diagramma indicato in figura: si tratta di un elemento dello spazio $C^0([-1,1])$.

Tenendo presente che i triangoli tratteggiati in figura sono uguali, e che l'area del rettangolo di base $[\sigma, 1]$ è uguale ad $(1-\sigma)$, ovviamente l'integrale di x_σ esteso all'intervallo $[-1,1]$ è uguale a $2(1-\sigma)$; siccome $x_\sigma(0) = -1$ e $\|x_\sigma\| = \max|x_\sigma| = 1$, si ha:

$$(2) \qquad \frac{|u_\sigma(x)|}{\|x_\sigma\|} = 2(1-\sigma) + 1 = 3 - 2\sigma.$$

Se ne deduce che non esistono maggioranti del primo membro della (1) minori di 3. Infatti qualunque sia $\varepsilon > 0$, esiste un elemento $u^{(\varepsilon)}$ di $C^0([-1,1])$ tale che il primo membro della (1) con $u = u^{(\varepsilon)}$ è maggiore di $3 - \varepsilon$: basta scegliere $\sigma < \varepsilon/2$ (oltre che minore di 1), porre $u^{(\varepsilon)} = u_\sigma$ e tener conto della (2).

Dunque l'estremo superiore del primo membro della (2) è 3, e pertanto:

$$\|u\| = 3.$$

2. Consideriamo il funzionale lineare:

$$u : x \in C^0([0,1]) \to \int_0^1 x(t)\,dt.$$

Qualunque sia $x \in C^0([0,1])$ si ha:

$$|u(x)| \leq \int_0^1 |x(t)|\,dt \leq \int_0^1 \|x\|\,dt = \|x\|,$$

e quindi u e' continuo, e risulta:

(3) $$\frac{|u(x)|}{\|x\|} \leq 1 \quad \forall x \in C^0([0,1]) - \{0\}.$$

Mostriamo che 1 e' l'estremo superiore del primo membro della (3). A tale scopo consideriamo la funzione $x_0 : t \in [0,1] \to 1$.

Siccome $|u(x_0)| = 1$, $\|x_0\| = \max |x_0| = 1$, esiste un elemento dell'insieme dei numeri a primo membro della (3), quello che si ottiene per $x = x_0$, che e' un maggiorante dell'insieme (perche' e' uguale ad 1). Dunque l'insieme e' dotato di massimo (= 1), e perciò risulta:

$$\|u\| = 1.$$

3. Consideriamo il funzionale lineare:

$$u : x \in C^1([0,3]) \to 8x(0) + 2x'(3).$$

Qualunque sia $x \in C^1([0,3])$ si ha:

$$|u(x)| \leq 8 \sup|x| + 2 \sup|x'| \leq 8 \sup|x| + 8 \sup|x'| = 8\,\|x\|$$

e quindi u e' continuo, e risulta:

(4) $$\frac{|u(x)|}{\|x\|} \leq 8.$$

Mostriamo che 8 e' l'estremo superiore del primo membro della (4). A tale scopo consideriamo la funzione $x_0 : t \in [0,3] \to (t-3)^2$.

Siccome $u(x_0) = 8 \cdot 9$ e $\|x_0\| = \max |x(t)| = x(0) = 9$, il primo membro della (4) per $x = x_0$ e' uguale ad 8. Con lo stesso discorso dell'esempio precedente si conclude che

$$\|u\| = 8.$$

7. Duale di uno spazio di Hilbert.

Nel numero 6 abbiamo introdotto il duale (forte) di uno spazio vettoriale normato: le considerazioni ivi svolte valgono in particolare nel caso di uno spazio pre-hilbertiano.

In questo numero vedremo che il duale di uno spazio di Hilbert è uno spazio di Hilbert ad esso equivalente , sicché ogni spazio di Hilbert si può identificare con il suo duale.

Ricordiamo intanto che, se *S* è uno spazio pre-hilbertiano , per ogni fissato $x \in S$ il funzionale definito in *S*:

(1) $$u_x : y \in S \to (x, y)_S$$

è lineare nel caso *S* reale , ma non nel caso *S* complesso , in cui è lineare il suo coniugato ([1]) : dunque in entrambi i casi u_x è un elemento del duale algebrico di *S* [pag.69]. Poichè il funzionale u_x è continuo ([2]), esso appartiene anche al duale forte *S** [pag.173] ; inoltre per ogni $x \in S$ risulta:

(1') $$\|u_x\|_{S^*} = \|x\| \quad .$$

DIM. (1'). Per la disuguaglianza di Schwarz si ha:

$$|u_x(y)| = \left|\overline{u_x(y)}\right| = |(y, x)| \le \|y\|\|x\| \quad \forall y \in S$$

e si ritrova così la continuità di u_x ; inoltre, supposto $x \neq 0$ [per *x*=0 il funzionale u_x è identicamente nullo e la (1') è banale], se ne deduce che:

$$\|u_x\|_{S^*} = \sup_{y \in S-\{0\}} \frac{|u_x(y)|}{\|y\|} \le \|x\|_S \, .$$

Da tale relazione segue subito la (1') , perchè sussiste la disuguaglianza opposta:

$$\|x\|_S = \frac{(x, x)}{\|x\|} = \frac{u_x(x)}{\|x\|} \le \sup_{y \in S-\{0\}} \frac{|u_x(y)|}{\|y\|} = \|u_x\|_{S^*} \, .$$

Orbene, se *S* è uno spazio di Hilbert, l'operatore di *S* in *S** :

(2) $$j : x \in S \to u_x \in S^* \ ,$$

ovviamente lineare, è invertibile ed ha per codominio tutto lo spazio *S** . Infatti sussiste il seguente **teorema di Riesz** :

(7.1) *Se S è uno spazio di Hilbert, allora per ogni* $u \in S^*$ *esiste uno ed un solo* $x \in S$ *tale che* $u = u_x$ *, ossia tale che*:

(3) $$< u, y >= (x, y) \quad \forall y \in S \ .$$

In altri termini, assegnato un funzionale lineare e continuo f definito nello spazio di Hilbert S, il problema di ricercare un elemento x di S per il quale risulti:

(3') $$(x, y) = < f, y > \quad \forall y \in S$$

ammette una e una sola soluzione.

([1]) Ciò perché il prodotto scalare è per definizione una forma sesquilineare .

([2]) La continuità è stata dimostrata [cfr. (4.2) di pag.97] in base alla definizione.

Nota.
La seconda formulazione della tesi appare superflua ma è molto importante, perchè utilissima in svariate questioni : ogni qualvolta un dato problema, ambientato in uno spazio di Hilbert, si riesce ad inquadrare in un'equazione del tipo $(3')$ nell'incognita $x \in S$, con termine noto $u \in S^*$, il teorema di Riesz fornisce per il problema un teorema di esistenza e unicità. Un esempio in proposito sarà riportato a pag.182.4 [esempio 10].

DIM. In primo luogo dimostriamo l'esistenza. Fissato $u \in S^*$ e supposto $u \neq 0$ ([3]) , esiste almeno un $x \in S$ tale che sia $u(x) \neq 0$, e quindi l'insieme chiuso:

(4) $$H_u = \{ y \in S : u(y) = 0 \}$$ ([4])

non coincide con S : essendo H_u (nucleo di u) un sottospazio chiuso di S diverso da S , per il teorema (4.6) di pag. 121 il suo supplemento ortogonale $H_u^{\perp}$ è diverso da $\{0\}$. Esiste allora almeno un elemento $x_u \in H_u^{\perp}$ diverso da 0: orbene, posto $e_u = \dfrac{x_u}{\|x_u\|}$, proveremo che il vettore di S proporzionale ad e_u

(5) $$x = u(e_u)e_u$$

verifica la (3).

Fissato $y \in S$ e detti $P_{H_u}, P_{H_u^{\perp}}$ gli operatori proiezione di S rispettivamente sui sottospazi supplementari H_u , $H_u^{\perp}$, per il teorema (4.4) di pag.119 si ha

(6) $$y = P_{H_u} y + P_{H_u^{\perp}} y \, .$$

Ricordiamo che H_u è un iperpiano di S [pag.63 ultimo comma, con $S_0 = H_u$ (ossia $x_0 = 0, \alpha = 0$] , e per la definizione di iperpiano (nel caso di un sottospazio) il suo supplemento ortogonale $H_u^{\perp}$ è una retta passante per l'origine [pag.52]: pertanto il sottospazio $H_u^{\perp}$ ha dimensione 1 ed è generato dal vettore e_u di norma unitaria. Essendo $H_u^{\perp}$ chiuso ([5]), per la (11) di pag.118 (con $H = H_u^{\perp}$ ed n=1) si ha allora :

$$P_{H_u^{\perp}} y = (y, e_u)e_u \ ,$$

e quindi per la (6) risulta ([6]):

(7) $$< u, y > = < u, P_{H_u} y > + < u, P_{H_u^{\perp}} y > = < u, P_{H_u^{\perp}} y > = < u, (y, e_u)e_u > \, .$$

D'altra parte l'ultimo membro della (7) si può esprimere come segue ([7]), tenendo presente la (5):

$$< u, (y, e_u)e_u > \; = (e_u, y) < u, e_u > = \; u(e_u)(e_u, y) \; = \; (u(e_u)e_u, y) = (x, y)$$

e dopo ciò dalla (7) consegue la (3).

Per dimostrare l'unicità, procedendo per assurdo ammettiamo che per un certo $u \in S^*$ esistano due punti distinti di S, x_1 ed x_2 , tali che per ogni $y \in S$ si abbia

$$< u, y > = (x_1, y) \quad , \quad < u, y > = (x_2, y)$$

e quindi

$$(x_1, y) = (x_2, y) \, , \quad \text{ossia} \quad (x_1 - x_2, y) = 0 \, .$$

Poiché tale relazione sussiste qualunque sia $y \in S$, in particolare per $y = x_1 - x_2$ ne segue che $(x_1 - x_2, x_1 - x_2) = 0$ ossia $\|x_1 - x_2\|^2 = 0$: assurdo in quanto $x_1 \neq x_2$. Dopo ciò il teorema è dimostrato.

([3]) Se $u(x)=0 \; \forall x \in S^*$, l'esistenza di un $x \in S$ tale da verificare la (3) è banale (basta porre x=0)

([4]) H_u è la controimmagine dell'insieme chiuso $\{0\}$ mediante la funzione $u \in S^*$ (definita in S ed ivi continua), e perciò è un insieme chiuso per la (7.4) di pag.22.

([5]) Il sottospazio $H_u^{\perp}$ è chiuso per la (1.9) di pag.109 [per la notazione ivi adoperata cfr.pag.96].

([6]) Si tenga presente che per la (4) si ha $u(P_{H_u} y) = 0$ [il vettore $P_{H_u} y$ appartiene banalmente ad H_u].

([7]) Essendo la dualità $<\cdot,\cdot>$ una forma sesquilineare, il fattore numerico $\lambda = (y, e_u)$ "trasportato fuori" diventa $\overline{\lambda} = (e_u, y)$.

Il teorema appena acquisito si può enunciare come segue:

(7.2) *Se S è uno spazio di Hilbert, l'operatore j definito dalla* (2) *è un isomorfismo isometrico dello spazio normato S sullo spazio normato S*.*

Infatti il teorema di Riesz asserisce che *j* è invertibile ed ha per codominio tutto lo spazio *S*.* ossia che *j* è un isomorfismo algebrico di *S* su *S**; peraltro si tratta di un isomorfismo isometrico perché per la (1′) risulta $\|jx\|_{S^*} = \|x\|$ per ogni $x \in S$.

L'operatore *j* definito dalla (2) si chiama **l'isomorfismo canonico** dello spazio di Hilbert *S* su *S** ; l'operatore inverso j^{-1} si chiama **l'isomorfismo canonico** di *S** su *S.*

Fissato $x \in S$ dalla (2) segue $jx = u_x$ e quindi per la definizione di u_x si ha:

(8) $$(x,y) = < jx, y > \qquad \forall y \in S \ ;$$

fissato $u \in S^*$ da tale uguaglianza si deduce poi la seguente, ponendo $j^{-1}u = x$ (sicché $jx=u$):

(9) $$< u, y > = (j^{-1}u, y) \qquad \forall y \in S \, .$$

Le (8),(9) mostrano che *se S è uno spazio di Hilbert, mediante l'isomorfismo canonico il prodotto scalare in S si trasforma nella dualità tra S* ed S, e viceversa quest'ultima si trasforma nel prodotto scalare in S.*

Il duale forte di uno spazio di Hilbert è uno spazio normato; mostriamo che si ha inoltre:

(7.3) *Il duale di uno spazio di Hilbert è a sua volta uno spazio di Hilbert.*

DIM. Si tratta di far vedere che è possibile definire nello spazio *S** un prodotto scalare *p*(*u*,v) in modo che per ogni $u \in S^*$ si abbia $p(u,u) = \|u\|^2_{S^*}$, e ciò si realizza subito utilizzando l'isomorfismo canonico di *S** su *S.* Infatti, posto

$$p(u,\mathrm{v}) = (j^{-1}u, j^{-1}\mathrm{v})_S \qquad \forall (u,\mathrm{v}) \in S^* \times S^* \, ,$$

per ogni $u \in S^*$ si ha:

$$p(u,u) = (j^{-1}u, j^{-1}u)_S = \left\| j^{-1}u \right\|^2_S = \|u\|^2_{S^*}$$

e l'asserto è provato.

Dopo ciò possiamo concludere, come preannunciato all'inizio di questo numero, che:

(7.4). *Il duale di uno spazio di Hilbert è uno spazio di Hilbert ad esso equivalente.*

8.Esempi.

1) Per la (7.4) il duale dello spazio di Hilbert $\mathbf{R}^k$ è equivalente ad $\mathbf{R}^k$, e perciò si può dire che *a meno di isomorfismi isometrici il duale di* $\mathbf{R}^k$ *coincide con* $\mathbf{R}^k$.

Ritroviamo così che tutti i funzionali lineari e continui definiti in $\mathbf{R}^k$ si ottengono ponendo, per ogni $c = (c_1, c_2, ..., c_k) \in \mathbf{R}^k$:

(10) $$< u_c , x > = c_1 x_1 + c_2 x_2 + ... + c_k x_k = (c,x)_{\mathbf{R}^k} \qquad \forall x = (x_1,...,x_k) \in \mathbf{R}^k ,$$

con l'aggiunta dell'uguaglianza:

$$\|u_c\|_{(\mathbf{R}^k)^*} = \|c\|_{\mathbf{R}^k} = \sqrt{c_1^2 + c_2^2 + ... + c_k^2} .$$

Quanto sopra vale anche se $S = \mathbf{C}^k$. Cogliamo l'occasione per una nota : siccome tutti gli elementi del duale sono definiti ponendo, per ogni $c = (c_1, c_2, ..., c_k) \in \mathbf{C}^k$:

(11) $$< u_c , x > = (c,x)_{\mathbf{C}^k} = c_1 \overline{x_1} + c_2 \overline{x_2} + ... + c_k \overline{x_k} \qquad \forall x = (x_1,...,x_k) \in \mathbf{C}^k$$

i funzionali del tipo (11) non sono lineari , ma è lineare il loro coniugato e perciò appartengono allo spazio duale di $\mathbf{C}^k$.

2).Qualunque spazio vettoriale Σ_k di dimensione finita *k* può essere munito, come andiamo a dimostrare, di prodotto scalare in modo da ottenere uno spazio di Hilbert. Ne dedurremo poi che Σ_k *è equivalente ad* $\mathbf{R}^k$ *se è sul campo reale, a* $\mathbf{C}^k$ *se è sul campo complesso*.

Supposto ad es. Σ_k sul campo reale, fissiamo una base di esso $(\mathbf{e}_1,...,\mathbf{e}_k)$. Per ogni $x \in \Sigma_k$ esiste ed è univocamente determinata una *k-pla* di numeri reali $\xi = (\xi_1,...,\xi_k)$, ossia un punto di $\mathbf{R}^k$, tale che sia

$$x = \xi_1 \mathbf{e}_1 + \xi_2 \mathbf{e}_2 + ... + \xi_k \mathbf{e}_k \ ;$$

analogamente qualunque punto $y \in \Sigma_k$ può essere espresso nella forma

$$y = \eta_1 \mathbf{e}_1 + \eta_2 \mathbf{e}_2 + ... + \eta_k \mathbf{e}_k$$

con $(\eta_1,...,\eta_k) = \eta \in \mathbf{R}^k$. Orbene, posto:

$$p(x,y) = (\xi,\eta)_{\mathbf{R}^k}$$

si definisce un prodotto scalare in Σ_k e quindi la coppia $S_k = (\Sigma_k, p)$ è uno spazio pre-hilbertiano. In conseguenza lo spazio normato S_k è isomorfo allo spazio normato $\mathbf{R}^k$ [cfr.(2.5) pag.141] : siccome quest'ultimo è completo, ne segue che anche S_k è completo e pertanto lo spazio S_k è uno spazio di Hilbert; è equivalente ad $\mathbf{R}^k$ perché si ha:

$$\|x\|_{S_k} = \sqrt{(x,x)_S} = \sqrt{p(\xi,\xi)} = \sqrt{(\xi,\xi)_{\mathbf{R}^k}} = \|\xi\|_{\mathbf{R}^k} .$$

Ne segue che anche il duale di S_k è equivalente ad $\mathbf{R}^k$, giacchè per il teorema (7.4) gli spazi S_k ed $(S_k)^*$ sono equivalenti.

Se lo spazio vettoriale Σ_k è sul campo complesso, si procede analogamente.

3). Per la (7.4) il duale dello spazio di Hilbert $L^2(\Omega)$, con Ω aperto di $\mathbf{R}^k$, è equivalente ad $L^2(\Omega)$, e perciò si può dire che *a meno di isomorfismi isometrici il duale di* $L^2(\Omega)$ *coincide con* $L^2(\Omega)$.

Dunque tutti i funzionali lineari e continui definiti in $L^2(\Omega)$ si ottengono ponendo

(12) $$< u_c , x > = \int_\Omega c(t)x(t)\,dt = (c,x)_{L^2(\Omega)} \qquad \forall x \in L^2(\Omega)$$

con $c \in L^2(\Omega)$, e risulta

$$\|u_c\|_{(L^2(\Omega))^*} = \|c\|_{L^2(\Omega)} .$$

4). Sia *p* un numero reale ≥ 1 , e sia p' tale che

$$\frac{1}{p} + \frac{1}{p'} = 1 \qquad (\ p' = +\infty \text{ se } p=1).$$

Per il duale di $L^p(\Omega)$, con Ω aperto di $\mathbf{R}^k$, se $p \neq 2$ non si può applicare il teorema di Riesz, dato che $L^p(\Omega)$ non è uno spazio di Hilbert (non è pre-hilbertiano). Tuttavia, come vedremo tra breve, valgono risultati analoghi a quelli del caso *p*=2 : precisamente per $p \geq 1$ (*ma non* per $p = +\infty$) il duale di $L^p(\Omega)$ è equivalente ad $L^{p'}(\Omega)$, e perciò si può dire che *a meno di isomorfismi isometrici il duale di* $L^p(\Omega)$

coincide con $L^{p'}(\Omega)$; invece il duale di $L^{\infty}(\Omega)$ non si può identificare con $L^1(\Omega)$: ce ne occuperemo nell'esempio seguente (per questo motivo abbiamo supposto $p \in [1,+\infty[$, escludendo così il duale di $L^{\infty}(\Omega)$).

Veniamo alla dimostrazione di quanto asserito.

Abbiamo visto [es.4 pag.146] che se $p \geq 1$, per ogni $c \in L^{p'}(\Omega)$ il funzionale lineare u_c definito dall'uguaglianza (12) $\forall x \in L^p(\Omega)$ è continuo [quindi è un elemento di $(L^p(\Omega))^*$], e che risulta:

$$|u_c(x)| \leq \|c\|_{L^{p'}(\Omega)} \|x\|_{L^p(\Omega)} \quad ;$$

ne segue che per $x \neq 0$ si ha:

$$\frac{|u_c(x)|}{\|x\|_{L^p(\Omega)}} \leq \| c \|_{L^{p'}(\Omega)} . \tag{13}$$

Orbene, un ulteriore teorema di Riesz assicura che qualunque sia $p \in [1,+\infty[$, per ogni funzionale lineare e continuo u definito in $L^p(\Omega)$ [cioè per ogni $u \in (L^p(\Omega))^*$] , esiste uno ed un solo $c \in L^{p'}(\Omega)$ tale che sia :

$$< u, x > \; = u_c(x) = \int_{\Omega} c(t)x(t)\,dt \qquad \forall x \in L^p(\Omega) \qquad : \tag{14}$$

inoltre risulta:

$$\|u\|_{(L^p(\Omega))^*} = \|u_c\|_{(L^p(\Omega))^*} = \|c\|_{L^{p'}(\Omega)} \ . \tag{15}$$

Cioè tutti i funzionali lineari e continui definiti in $L^p(\Omega)$ sono del tipo (14), proprio come nel caso p=2 in cui $L^p(\Omega)$ è uno spazio di Hilbert ; però, a differenza del caso p=2, l'integrale nella (14) non è un prodotto scalare : basta osservare che i fattori c ed x non appartendono al medesimo spazio.

Rinunciando alla dimostrazione dell'esistenza e unicità del predetto $c \in L^{p'}(\Omega)$ (non è affatto semplice), ci limitiamo a dimostrare la (15). Siccome la norma di u_c è l'estremo superiore al variare di x in $L^{p'}(\Omega) - \{0\}$ del rapporto a primo membro della (13), per la stessa (13) tale norma è $\leq \|c\|_{L^{p'}(\Omega)}$: basta allora dimostrare che qualunque sia $c \in L^{p'}(\Omega) - \{0\}$ esiste $x_0 \in L^{p'}(\Omega) - \{0\}$ tale che ([8])

$$\frac{|u_c(x_0)|}{\|x_0\|_{L^p(\Omega)}} = \|c\|_{L^{p'}\Omega)} . \tag{16}$$

A tale scopo consideriamo la funzione x_0 definita quasi ovunque in Ω ponendo $x_0(t) = |c(t)|^{p'-2} c(t)$; essendo $c \in L^{p'}(\Omega)$ e

$$|x_0(t)|^p = |c(t)|^{(p'-1)p} = |c(t)|^{p'} ,$$

la funzione x_0 è un elemento di $L^p(\Omega)$, e risulta:

$$\|x_0\|_{L^p(\Omega)} = (\int_{\Omega} |c(t)|^{p'})^{\frac{1}{p}} = \|c\|_{L^{p'}\Omega)}^{\frac{p'}{p}} .$$

D'altra parte si ha $c(t)x_0(t) = |c(t)|^{p'}$ e quindi per la (14)

$$u_c(x_0) = \int_{\Omega} c(t)x_0(t)\,dt = \int_{\Omega} |c(t)|^{p'}\,dt = \|c\|_{L^{p'}\Omega)}^{p'} ;$$

ne segue che risulta:

$$\frac{|u_c(x_0)|}{\|x_0\|_{L^p(\Omega)}} = \frac{\|c\|_{L^{p'}\Omega)}^{p'}}{\|c\|_{L^{p'}\Omega)}^{\frac{p'}{p}}} = \|c\|_{L^{p'}\Omega)}$$

e pertanto x_0 verifica la (16).

([8]) La (15) è banale per c=0, dato che per la (14) c=0 $\Rightarrow$ u_c=0.

La (15) è dunque acquisita; allora l'operatore lineare che ad ogni $u \in (L^p(\Omega))^*$ associa l'unico elemento *c* di $L^{p'}(\Omega)$ per il quale sussiste la (14), è un isomorfismo isometrico di $(L^p(\Omega))^*$ su $L^{p'}(\Omega)$, e pertanto i due spazi sono equivalenti.

5). Per quanto si è detto nell'esempio 4), qualunque siano i numeri reali p, p' maggiori di 1 e tali che $1/p + 1/p' = 1$, ciascuno degli spazi $L^p(\Omega)$, $L^{p'}(\Omega)$ si identifica con il duale dell'altro (del resto abbiamo già visto a pag.188 che per $p \in]1,+\infty[$ lo spazio $L^p(\Omega)$ è riflessivo); inoltre il duale di $L^1(\Omega)$ si identifica con $L^\infty(\Omega)$.

Invece il duale di $L^\infty(\Omega)$ non si identifica con $L^1(\Omega)$, ma contiene strettamente uno spazio normato *H* equivalente ad $L^1(\Omega)$. Perciò si può dire che *a meno di isomorfismi isometrici il duale di* $L^\infty(\Omega)$ *contiene strettamente* $L^1(\Omega)$.

Allo scopo di dimostrare ciò, osserviamo intanto che un elemento del duale di $L^\infty(\Omega)$ si ottiene, come nel caso $p \in [1,+\infty[$, fissando un elemento *c* di $L^1(\Omega)$ e considerando il funzionale u_c definito ponendo

$$u_c(x) = \int_\Omega c(t)x(t)\,dt \qquad \forall x \in L^\infty(\Omega). \tag{17}$$

La continuità di u_c discende dalle relazioni seguenti:

$$|< u_c, x >| = \leq \int_\Omega |c(t)||x(t)|dt \leq \sup_{t\in\Omega} |x(t)| \int_\Omega |c(t)|\,dt \leq \|x\|_{L^\infty(\Omega)} \|c\|_{L^1(\Omega)} \qquad \forall x \in L^\infty(\Omega),$$

dalle quali consegue poi

$$\|u_c\|_{(L^\infty(\Omega))^*} = \sup_{x\in L^\infty(\Omega)} \frac{|u_c(x)|}{\|x\|_{L^\infty(\Omega)}} \leq \|c\|_{L^1(\Omega)}\ . \tag{18}$$

Si vede subito che nella (18) vale il segno di uguaglianza. Basta osservare che, ponendo $x_0(t) = \text{signum}c(t)$, si ha $x_0 \in L^\infty(\Omega)$, $\|x_0\|_{L^\infty(\Omega)} = 1$ e quindi tenendo presente la (17):

$$\frac{u_c(x_0)}{\|x_0\|_{L^\infty(\Omega)}} = u_c(x_0) = \int_\Omega c(t)x_0(t)\,dt = \int_\Omega |c(t)|\,dt = \|c\|_{L^1(\Omega)}.$$

Pertanto la (18) diviene:

$$\|u_c\|_{(L^\infty(\Omega))^*} = \|c\|_{L^1(\Omega)}, \tag{19}$$

sicché l'operatore lineare *A* che ad ogni $c \in L^1(\Omega)$ associa $u_c \in (L^\infty(\Omega))^*$ è un isomorfismo isometrico di $L^1(\Omega)$ sul codominio di *A*, diciamolo *H*, che è un sottospazio di $(L^\infty(\Omega))^*$; per la (19) si ha che *H* è equivalente ad $L^1(\Omega)$. Siccome $(L^\infty(\Omega))^* \supseteq H$, con tale identificazione si può dire che $(L^\infty(\Omega))^*$ contiene $L^1(\Omega)$.

Orbene, si dimostra che esistono funzionali lineari e continui definiti in $L^\infty(\Omega)$ che non sono del tipo di u_c definito dalla (17), ossia che esistono elementi di $(L^\infty(\Omega))^*$ che non appartengono ad *H*. Pertanto si può dire che il duale di $L^\infty(\Omega)$ contiene strettamente $L^1(\Omega)$.

6). Per costruire un elemento del duale di $H^{m,p}(\Omega)$ con $p \in]1,+\infty[$, denotiamo con *I* l'insieme dei multiindici $\alpha = (\alpha_1,...,\alpha_k)$ tali che $|\alpha| \leq m$, e fissiamo una famiglia $c = (c_\alpha)_{\alpha\in I} \in H^{m,p'}(\Omega)$, con p' al solito tale che $1/p + 1/p' = 1$. Mostriamo che il funzionale lineare u_c definito ponendo:

(20) $$< u_c , x >= \sum_{|\alpha|=0}^{m} \int_{\Omega} c_\alpha(t) D^\alpha x(t) dt \quad \forall x \in H^{m.p}(\Omega)$$

è continuo, e quindi appartiene ($(H^{m.p}(\Omega))^*$.

Infatti, tenendo presente la disuguaglianza di Schwarz-Hölder [cfr.(3) pag.54] si ha:

$$|u_c(x)| \le \sum_{|\alpha|=0}^{m} \int |c_\alpha(t)| |D^\alpha x(t)| dt \le \sum_{|\alpha|=0}^{m} \|c_\alpha\|_{L^{p'}(\Omega)} \|D^\alpha x\|_{L^p(\Omega)} \le \|x\|_{H^{m,p}(\Omega)} \sum_{|\alpha|=0}^{m} \|c_\alpha\|_{L^{p'}(\Omega)} .$$

Orbene, si dimostra che tutti gli elementi di $(H^{m.p}(\Omega))^*$ sono del tipo (20); precisamente, per ogni $u \in (H^{m.p}(\Omega))^*$ esiste almeno una famiglia $c=(c_\alpha)_{\alpha \in I} \in H^{m.p'}(\Omega)$ tale che si abbia $u = u_c$; però si prova che la famiglia c non è univocamente determinata da u , sicché le norma di u in $(H^{m.p}(\Omega))^*$ non può essere uguale alla norma di c in $H^{m.p'}(\Omega)$, cioè al numero

$$\nu(c) = \left(\sum_{|\alpha|=0}^{m} \|c_\alpha\|_{L^{p'}\Omega)}^{p'} \right)^{\frac{1}{p'}}$$

che non dipende soltanto da *u*, ma anche da *c*. Peraltro, detto C_u l'insieme costituito dalle famiglie $c= (c_\alpha)_{\alpha \in I} \in H^{m.p'}(\Omega)$ tali che $u = u_c$, si dimostra che risulta:

$$\|u\|_{(H^{m.p}(\Omega))^*} = \inf_{c \in C_u} \nu(c) .$$

Dopo ciò si può dire che il duale di $H^{m.p'}(\Omega)$ si identifica con $H^{m.p'}(\Omega)$. Tale conclusione vale anche se allo spazio $H^{m,p}(\Omega)$ subentra uno degli spazi $H_0^{m,p}(\Omega)$, $W^{m,p}(\Omega)$.

7) Fissato $c \in L^1(\Omega)$, un elemento del duale di $C^0(\Omega)$ è il funzionale lineare

$$u_c : x \in C^0(\Omega) \to \int_{\Omega} c(t) x(t)\, dt ,$$

che è continuo perché risulta:

$$|u_c(x)| \le \int_{\Omega} |c(t)| |x(t)|\, dt \le \sup_{t \in \Omega} |x(t)| \int_{\Omega} |c(t)|\, dt = \|x\|_{C^0(\Omega)} \|c\|_{L^1(\Omega)} .$$

Un elemento del duale di $C^0(\Omega)$ si ottiene anche considerando lo spazio $L^1(\Omega, \mu)$, dove μ è una misura nell'insieme delle parti di Ω misurabili secondo Lebesgue. Infatti fissato $c \in L^1(\Omega, \mu)$, il funzionale lineare

$$u_c : x \in C^0(\Omega) \to \int_{\Omega} cx\, d\mu$$

è continuo perché risulta:

$$|u_c(x)| \le \int_{\Omega} |c| |x|\, d\mu \le \sup_{t \in \Omega} |x(t)| \int_{\Omega} |c|\, d\mu = \|x\|_{C^0(\Omega)} \|c\|_{L^1(\Omega, \mu)} .$$

8) Il duale topologico di $C_0^\infty(\Omega)$ è lo spazio vettoriale costituito dalle distribuzioni su Ω : tale asserzione è banale, dato che discende direttamente dalle definizioni di duale topologico [pag.171] e di distribuzione [pag.151]; abbiamo voluto evidenziarla perché talvolta le distribuzioni vengono definite proprio come gli elementi del duale topologico di $C_0^\infty(\Omega)$, ossia , con il simbolo da noi adottato per il duale topologico, come gli elementi di $\mathcal{D}^*(\Omega)$ [abbiamo già detto che lo spazio $C_0^\infty(\Omega)$ si denota anche con $\mathcal{D}(\Omega)$].

Aggiungiamo una nota: poiché ogni elemento $z \in L^1_{loc}(\Omega)$ si identifica con la distribuzione T_z [pag.152], con tale identificazione si può dire che $L^1_{loc}(\Omega)$ è un sottospazio vettoriale di $\mathcal{D}^*(\Omega)$.

9) Per indicare un'applicazione del teorema di Riesz, consideriamo il problema seguente , detto problema omogeneo di Dirichlet **generalizzato**, oppure **in forma debole**.

Assegnata una funzione $f \in L^2(\Omega)$, con Ω aperto limitato di $\mathbf{R}^2$, determinare $u \in H_0^1(\Omega)$ tale che si abbia:

(21) $$\Delta_2 u = -f \quad \textit{nel senso delle distribuzioni su } \Omega .$$

Prima di dimostrare che tale problema ammette una e una sola soluzione, che si chiama **soluzione debole** del problema omogeneo di Dirichlet, occorre specificare il significato della (21).

L'operatore di Laplace

$$\Delta_2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} ,$$

inteso nel senso delle distribuzioni su Ω , significa che le due derivate parziali sono espresse dalle relazioni [cfr. (46) pag.158.2]:

(22) $$< \frac{\partial^2 u}{\partial x^2}, \varphi > = \iint_\Omega u \frac{\partial^2 \varphi}{\partial x^2} dxdy , \quad < \frac{\partial^2 u}{\partial y^2}, \varphi > = \iint_\Omega u \frac{\partial^2 \varphi}{\partial y^2} dxdy \quad \forall \varphi \in C_0^\infty(\Omega) \ .$$

D'altra parte la *u*, essendo elemento di $H_0^1(\Omega)$, è dotata di derivate parziali prime in senso debole; queste per definizione sono gli elementi di $L^2(\Omega)$ tali che si abbia [cfr.(6) pag.61]:

$$\iint_\Omega \frac{\partial u}{\partial x} \psi \, dxdy = -\iint_\Omega u \frac{\partial \psi}{\partial x} dxdy \ , \quad \iint_\Omega \frac{\partial u}{\partial y} \psi \, dxdy = -\iint_\Omega u \frac{\partial \psi}{\partial y} dxdy \quad \forall \psi \in C_0^\infty(\Omega) .$$

Applicando queste uguaglianze rispettivamente per $\psi = \frac{\partial \varphi}{\partial x}$ e per $\psi = \frac{\partial \varphi}{\partial y}$, si ottengono le seguenti:

$$\iint_\Omega \frac{\partial u}{\partial x} \frac{\partial \varphi}{\partial x} dxdy = -\iint_\Omega u \frac{\partial^2 \varphi}{\partial x^2} dxdy \quad , \quad \iint_\Omega \frac{\partial u}{\partial y} \frac{\partial \varphi}{\partial y} dxdy = -\iint_\Omega u \frac{\partial^2 \varphi}{\partial y^2} dxdy \quad \forall \varphi \in C_0^\infty(\Omega) ,$$

e quindi le (22) divengono rispettivamente

$$< \frac{\partial^2 u}{\partial x^2}, \varphi > = -\iint_\Omega \frac{\partial u}{\partial x} \frac{\partial \varphi}{\partial x} dxdy , \qquad < \frac{\partial^2 u}{\partial y^2}, \varphi > = -\iint_\Omega \frac{\partial u}{\partial y} \frac{\partial \varphi}{\partial y} dxdy \ .$$

Pertanto si ha :

$$< \frac{\partial^2 u}{\partial x^2}, \varphi > + < \frac{\partial^2 u}{\partial y^2}, \varphi > = -\iint_\Omega [\frac{\partial u}{\partial x} \frac{\partial \varphi}{\partial x} + \frac{\partial u}{\partial y} \frac{\partial \varphi}{\partial y}] dxdy = -\iint_\Omega \nabla u \bullet \nabla \varphi \ dxdy = -(\nabla u, \nabla \varphi_n)_{L^2(\Omega)}$$

e perciò $\Delta_2 u$, ossia il primo membro della (21), è la distribuzione definita come segue:

(23) $$< \Delta_2 u, \varphi > \ = \ -(\nabla u, \nabla \varphi)_{L^2(\Omega)} \qquad \forall \varphi \in C_0^\infty(\Omega) .$$

Naturalmente anche la funzione *f* a secondo membro della (21) va intesa nel senso delle distribuzioni ([9]), cioè identificata con la distribuzione $-T_f$ che è il funzionale definito dalla relazione [cfr.(20) pag.151]:

(24) $$- < T_f, \varphi > \ = \ -\iint_\Omega f \varphi \, dxdy = \ -(f, \varphi)_{L^2(\Omega)} \quad \forall \varphi \in C_0^\infty(\Omega) \quad ([10]).$$

In definitiva l'equazione (21) [nell'incognita *u*] si specifica come segue:

(25) $$(\nabla u, \nabla \varphi)_{L^2(\Omega)} = (f, \varphi)_{L^2(\Omega)} \quad \forall \varphi \in C_0^\infty(\Omega) .$$

([9]) Cfr.pag.152. Occorre tener presente che f , elemento di $L^2(\Omega)$, appartiene anche ad $L^1_{loc}(\Omega)$ [cfr. quanto osservato all' inizio di pag. 59].

([10]) Per esprimere l'integrale come prodotto scalare, l'elemento φ di $C_0^\infty(\Omega)$ va identificato con l'elemento di $L^2(\Omega)$ (classe di funzioni a due a due uguali q.o.) a cui appartiene.

Per dimostrare l'esistenza di una soluzione, osserviamo che, essendo $C_0^\infty(\Omega)$ denso in $H_0^1(\Omega)$ ([11]), per ogni $v \in H_0^1(\Omega)$ esiste una successione $(\varphi_n)_{n\in\mathbf{N}}$ di elementi di $C_0^\infty(\Omega)$ tale che:

(26) $$\lim_{n\to\infty} \varphi_n = v \quad \text{in } H_0^1(\Omega),$$

e questa comporta che

(27) $$\lim_{n\to\infty} \nabla\varphi_n = \nabla\mathbf{v} \quad \text{in } L^2(\Omega).$$

Allora, se esiste una soluzione u della (25), il primo membro della (25) con φ_n in luogo di φ, ossia $(\nabla u, \nabla\varphi_n)_{L^2(\Omega)}$, per la (27) tende a $(\nabla u, \nabla\mathbf{v})_{L^2(\Omega)}$, e questo numero è il prodotto scalare di u e $\mathbf{v}$ nello spazio $H_0^1(\Omega)$ [cfr. la seconda delle (2) di pag.104].

D'altra parte, siccome per la (26) risulta $\lim_{n\to\infty} \varphi_n = v$ anche nello spazio $L^2(\Omega)$, il secondo membro della (25) con φ_n in luogo di φ, ossia $(f, \varphi_n)_{L^2(\Omega)}$, tende all'integrale

(28) $$\iint_\Omega f\, v\, dxdy,$$

che si può scrivere nella forma $< f, \mathbf{v} >$, identificando f con il funzionale che ad ogni $v \in H_0^1(\Omega)$ associa l'integrale (28).

Pertanto dalla (25) scritta con φ_n in luogo di φ, passando al limite per $n \to \infty$ si deduce che un'eventuale soluzione u della (25) deve verificare l'uguaglianza seguente:

(29) $$(u, v)_{H_0^1(\Omega)} = < f, v > \qquad \forall v \in H_0^1(\Omega).$$

Riguardando la (29) come equazione nell'incognita $u \in H_0^1(\Omega)$, per il teorema di Riesz [si veda la (3') di pag.177] essa ammette una e una sola soluzione : orbene, si prova subito che tale soluzione u verifica la (25), ossia è soluzione della (21). Basta osservare che qualunque sia $\varphi \in C_0^\infty(\Omega)$, dalla (29) per $v=\varphi$ ([12]) si ottiene l'uguaglianza:

(30) $$(u, \varphi)_{H_0^1(\Omega)} = (f, \varphi)_{L^2(\Omega)} \qquad \forall \varphi \in C_0^\infty(\Omega)$$

che è appunto la (25) ([13]).

L'unicità per la (25) è ovvia, avendo dianzi stabilito che un'eventuale soluzione della (25) deve verificare la (29) e questa, come già detto, ha una sola soluzione.

10. Un'applicazione del problema omogeneo di Dirichlet in forma debole .

Consideriamo il seguente problema: assegnata una distribuzione di carica r (chiamata *densità di carica*) occupante una regione limitata del piano, calcolare il potenziale elettrico u da essa generato nell'interno della regione, determinato dalla condizione che sia nullo sulla frontiera.

Il gradiente di u cambiato di segno è il campo elettrico E, che è legato ad r dalla relazione

$$\operatorname{div} E = 4\pi r/\varepsilon,$$

dove ε è la costante dielettrica del mezzo. Dall'uguaglianza

$$-\nabla u = E, \quad \text{ossia} \quad \nabla u = -E,$$

si deduce

$$\operatorname{div}\nabla u = \operatorname{div}(-E) = -4\pi r/\varepsilon \quad \text{ossia} \quad \Delta_2 u = -4\pi r/\varepsilon.$$

([11]) Cfr. pag.104, terzo comma. Naturalmente ciascun elemento di $C_0^\infty(\Omega)$ va identificato con l'elemento di $H_0^1(\Omega)$ (classe di funzioni a due a due uguali q.o.) a cui appartiene.

([12]) Si veda quanto osservato nella precedente nota ([11]).

([13]) Per quanto attiene al primo membro della (30), si tenga presente il prodotto scalare nello spazio $H_0^1(\Omega)$, già richiamato poc'anzi; per il secondo membro della (30), si osservi la (24) identificando nel primo membro T_f con f.

Pertanto, posto $f = 4\pi\, r/\, \varepsilon$ e detto Ω l'interno della regione, la funzione u è soluzione nell'aperto Ω dell'equazione $\Delta_2 u = - f$ (*equazione di Poisson*), e fuori di Ω (dove $r = 0$) soluzione dell'equazione $\Delta_2 u = 0$ (*equazione di Laplace*); per la condizione al contorno u=0 su $\partial\Omega$, la funzione u è soluzione del problema omogeneo di Dirichlet per l'equazione di Poisson:

$$\begin{cases} \Delta_2 u = -f & in\ \Omega \\ u = 0 & \text{su } \partial\Omega \end{cases} .$$

Il problema in senso classico si formula come segue: assegnata una funzione f continua in un aperto Ω , ricercare una funzione u continua in $\Omega \cup \partial\,\Omega$, nulla su $\partial\Omega$, dotata di derivate prime e seconde continue in Ω ed ivi avente laplaciano uguale a $-f$. Se Ω non soddisfa ad opportune condizioni (ad es. che abbia frontiera "*di classe* C^2 "), non è detto che il problema ammetta soluzioni , anche se si suppone che il dato f sia una funzione continua non solo in Ω ma pure sulla frontiera di Ω. Lo spazio ambiente del problema è $C^2(\Omega) \cap C^0(\overline{\Omega})$.

Come appena evidenziato, per dimostrare l'esistenza e l'unicità della soluzione occorrono delle ipotesi su Ω, che sono troppo restrittive per le applicazioni. Per ottenere l'esistenza in ipotesi più generali, non solo su Ω ma anche sulla funzione f, occorre ambientare il problema in uno spazio "più ampio" : perciò si considera il problema omogeneo di Dirichlet generalizzato introdotto nell'esempio precedente che, come si è visto in base al teorema di Riesz, ammette nello spazio $H_0^1(\Omega)$ una e una sola soluzione.

E' ben noto che l'equazione di Poisson e quella di Laplace intervengono frequentemente nelle scienze applicate. Per esempio, la temperatura di un corpo omogeneo ed isotropo in condizioni di equilibrio è una funzione armonica (Δ_2 u = 0) ; in questo senso l'equazione di Laplace descrive il caso stazionario (cioè indipendente dal tempo) dell'equazione di diffusione. La configurazione di equilibrio di una membrana perfettamente elastica è una funzione armonica, come pure il potenziale di velocità di un fluido omogeneo, o il potenziale di un campo attorno ad un profilo alare.

L'equazione di Poisson svolge un ruolo importante nella teoria dei campi conservativi (campo elettrico, magnetico, gravitazionale, ecc.) , cioè nel caso che il vettore campo è esprimibile come il gradiente di un potenziale.

In generale le formulazioni deboli di problemi di questo tipo , oltre ad assicurare l'esistenza di soluzioni (*deboli*), comportano diversi vantaggi e in particolare conducono in maniera naturale allo sviluppo di metodi numerici *di tipo Galerkin.*

Note sui modelli matematici.

L'esempio dianzi esposto ci offre l'occasione per effettuare un discorso di carattere generale.

I teoremi di esistenza e unicità non interessano solo i matematici.

Quando si traduce un problema fisico in un problema matematico, si vengono a scrivere delle equazioni e/o disequazioni, nonché delle uguaglianze e/o disuguaglianze , che nel loro complesso costituiscono quello che si chiama un *modello matematico* del problema fisico considerato; è importante che il problema possa essere inquadrato in una equazione nell'incognita x del tipo:

(31) $$Ax=y \ ,$$

dove $A: S \to T$ ed y rappresentano i dati del problema, e la soluzione si cerca nell'insieme S .

Nella costruzione di un modello matematico può accadere che le considerazioni fisiche atte ad individuare la soluzione non si facciano intervenire *tutte* (inconsciamente), oppure che si inseriscano condizioni superflue; inoltre spesso si effettuano delle semplificazioni, o si trascura "qualcosa" per ragioni di semplicità. A causa di ciò, non è detto che il modello costruito sia "buono" , cioè che abbia le potenzialità per rappresentare efficacemente il fenomeno da studiare.

Per essere sicuri che il modello sia "valido ", occorre che il problema matematico sia come suol dirsi **ben posto**: ciò significa che debbono essere verificate due condizioni.

a) *Il problema deve ammettere una e una sola soluzione.*

L'esistenza garantisce che nel costruire il modello matematico non si è richiesto più del necessario, l'unicità garantisce che non si è omesso qualcosa di essenziale; con riferimento alla (31), l'esistenza significa che l'insieme T è il codominio dell'operatore A (cioè che la corrispondenza $A: S \to T$ è suriettiva), l'unicità significa che A è invertibile (cioè che la corrispondenza è iniettiva , o biunivoca).

b) *La soluzione deve dipendere con continuità dai dati*: ossia, con riferimento alla (31), *la funzione inversa di* $A: S \to T$ *deve essere continua.*

Adoperando un linguaggio intuitivo, la **b)** richiede che se si approssimano i dati del problema [come accade spesso quando si inseriscono i dati nel computer per il calcolo della soluzione] si ottenga effettivamente una soluzione approssimata ; ossia impone che un *piccolo* errore sui dati provochi un errore *piccolo* sulla soluzione.

In termini precisi, la **b)** richiede che la corrispondenza *dati* $\to$ *soluzione* , appunto la funzione inversa di $A: S \to T$, sia una funzione continua. Naturalmente per parlare di continuità bisogna che S e T siano spazi topologici, e la continuità è legata alle topologie dei due spazi; anche il significato della parola *approssimazione* dipende dalle due topologie. E' chiaro allora che le topologie per S e per T (quindi i due spazi topologici), vanno scelte in modo adeguato alla questione da studiare.

Se S e T sono spazi di Banach, e se la corrispondenza A è un operatore lineare, dopo aver accertato che A è invertibile e suriettivo, per constatare che l'operatore inverso è continuo basta stabilire che A è continuo [teorema di Banach, (2.4) pag.141].

* * *

Per concludere, mostriamo che il problema espresso dalla (21) di pag.182.2 è ben posto. Avendo già stabilito la condizione **a),** si tratta di provare la **b).**

Detto A l'operatore lineare che ad ogni $u \in H_0^1(\Omega)$ associa il funzionale definito come segue:

$$\mathrm{v} \in H_0^1(\Omega) \to (u,\mathrm{v})\, H_0^1(\Omega)$$

elemento del duale di $H_0^1(\Omega)$, l'equazione (21) si esprime nella forma (31) con $y = -f$, dove f è un elemento di $L^2(\Omega)$ identificato con un elemento di $(H_0^1(\Omega))^*$ [cfr. pag.182.3 dopo la (28)].

Siccome A è un isomorfismo dello spazio di Hilbert $H_0^1(\Omega)$ sul duale di esso (quindi tra spazi di Banach), l'operatore inverso è continuo per il teorema di Banach dianzi citato.

11. Come ulteriore applicazione del teorema di Riesz dimostriamo che ([14]):

(8.1) *Se* Ω *è un aperto limitato di* **R** , *ogni funzione* $u \in L^2(\Omega)$ *è dotata di primitiva debole, anch'essa elemento di* $L^2(\Omega)$: *esiste cioè* $w \in L^2(\Omega)$ *tale che*

$$w' = u \ .$$

DIM. Supposto $H_0^1(\Omega)$ munito della norma (2) di pag. 104, che è equivalente a quella usuale in quanto Ω è limitato, consideriamo il funzionale lineare:

$$f: \quad \mathrm{v} \in H_0^1(\Omega) \to \int_\Omega u(t)\mathrm{v}(t)dt$$

([14]) L'osservazione è stata fatta dall'ing. MAURO RUSSO durante una mia lezione del corso di Dottorato in Ingegneria Industriale. Con tale citazione intendo manifestare all'ing. Russo il mio apprezzamento e la mia stima.

che è continuo [es.10 pag. 148]. Per il teorema di Riesz [(7.1) pag. 177] esiste allora $z \in H_0^1(\Omega)$ tale che:

$$(z, \mathrm{v})_{H_0^1(\Omega)} = < f, \mathrm{v} > = \int_\Omega u(t)\mathrm{v}(t)dt \qquad \forall\, \mathrm{v} \in H_0^1(\Omega)$$

In particolare per $\mathrm{v} = \varphi \in C_0^\infty(\Omega)$ [inteso come elemento di $H_0^1(\Omega)$], tenuto conto della seconda delle (2) di pag.104 con k=1, dalla precedente relazione si deduce:

$$\int_\Omega z'(t)\varphi'(t)dt = \int_\Omega u(t)\varphi(t)dt \quad \forall\ \varphi \in C_0^\infty(\Omega). \tag{32}$$

Orbene, posto:

$$w = -z',$$

si ha $w \in L^2(\Omega)$ e la (32) si scrive:

$$\int_\Omega u(t)\varphi(t)dt = -\int_\Omega w(t)\varphi'(t)dt \quad \forall\ \varphi \in C_0^\infty(\Omega).$$

Dunque, per la definizione di derivata debole [(6) pag.61, con lo scambio delle lettere u e w] , l'elemento u di $L^2(\Omega)$ è la derivata debole di w e l'asserto è provato.

9. Il teorema di Hahn-Banach. Alcuni corollari.

Fondamentale è il seguente teorema di Hahn-Banach, del quale non riportiamo la dimostrazione:

(9.1). Siano: S uno spazio vettoriale normato, H un sottospazio di S.

Per ogni funzionale lineare u_0, definito e continuo in H, esiste almeno un prolungamento u di u_0 su S, lineare, continuo e tale che:

$$\|u\|_{S^*} = \|u_0\|_{H^*} .$$

Dal teorema enunciato sono stati dedotti numerosi corollari, non meno importanti del teorema stesso. Riportiamo i seguenti.

(9.2) Siano: S uno spazio vettoriale normato, H un sottospazio di S.

Se H non è denso in S, esistono funzionali lineari, definiti e continui in S, che sono identicamente nulli in $\overline{H}$ ma non in tutto S.

In altri termini: per ogni $x_0 \notin \overline{H}$ esiste almeno un funzionale lineare u, definito e continuo in S, tale che:

(1) $$u(x_0) = 1 \quad , \quad u(x) = 0 \quad \forall x \in \overline{H} .$$

— — — — — — — — —

Dim. Poichè $x_0 \notin \overline{H}$, il sottospazio di S:

$$H_0 = \overline{H} + K x_0$$

è somma diretta dei due sottospazi $\overline{H}$, Kx_0 [cfr. pag. 52], e quindi per ogni $x \in H_0$ esistono $h \in \overline{H}$ e $\lambda \in K$, univocamente determinati, tali che:

(2) $$x = h + \lambda x_0 .$$

Denotiamo allora con u_0 il funzionale che ad ogni $x \in H_0$ associa il numero λ per il quale sussiste la (2). Ovviamente u_0 è lineare, associa il numero 1 al punto x_0, ed è identicamente nullo in $\overline{H}$: ossia, le (1) sono verificate dal funzionale u_0.

D'altra parte, u_0 è anche continuo. Infatti qualunque sia il punto (2), supposto non appartenente ad $\overline{H}$ (ricchè $\lambda \neq 0$), si ha:

$$\frac{|u_0(x)|}{\|x\|} = \frac{|\lambda|}{\|h + \lambda x_0\|} = \frac{1}{\left\|\frac{1}{\lambda} h + x_0\right\|} = \frac{1}{\|x_0 - h'\|} \leq \frac{1}{\delta}$$

dove $h' = -h/\lambda$ e δ denota la distanza di x_0 ed $\overline{H}$, che è positiva; ne segue che:

$$|u_0(x)| \leq \frac{1}{\delta} \|x\| \quad \forall x \in H_0 \qquad (1)$$

donde la continuità di u_0.

Per il teorema di Hahn-Banach (con H_0 in luogo di H), esiste allora un prolungamento u di u_0 su S, lineare e continuo. Siccome u verifica le (1), l'asserto è dimostrato.

— — — — — — — —

(1) La disuguaglianza, stabilita per $x \in H_0 - \overline{H}$, è verificata banalmente se $x \in \overline{H}$ in quanto u_0 è identicamente nullo in $\overline{H}$.

(9.3) Per ogni punto $x \neq 0$ dello spazio vettoriale normato S, esiste almeno un elemento non nullo u del duale S^* tale che:

$$|\langle u, x\rangle| = \|u\| \cdot \|x\| , \tag{3}$$

e quindi anche un elemento u_1 di S^*, avente norma uguale ad 1, tale che:

$$|\langle u_1, x\rangle| = \|x\| . \tag{4}$$

Osserviamo che, essendo:

$$|\langle u, x\rangle| \leq \|u\| \cdot \|x\| \qquad \forall u \in S^* \text{ e } \forall x \in S ,$$

e quindi:

$$\sup_{u \in S^* - \{0\}} \frac{|\langle u, x\rangle|}{\|u\|} \leq \|x\| \qquad \forall x \in S , \tag{5}$$

in virtù del teorema (9.3) per ogni $x \in S - \{0\}$ l'estremo superiore (5) è un massimo ed è uguale a $\|x\|$. Si ha dunque:

$$\sup_{u \in S^* - \{0\}} \frac{|\langle u, x\rangle|}{\|u\|} = \max_{u \in S^* - \{0\}} \frac{|\langle u, x\rangle|}{\|u\|} = \|x\| \qquad \forall x \in S . \tag{6}$$

Dim. (9.3). Fissato un punto $x_0 \in S - \{0\}$, sia H il sottospazio di S generato da x_0, cioè il sottospazio costituito dai vettori di S proporzionali ad x_0. Sia quindi u_0 il funzionale che ad ogni $x \in H$ associa il numero λ tale che $x = \lambda x_0$, e sia $u_0' = \|x_0\| \cdot u_0$.

Ovviamente u_0, quindi anche u_0', è lineare; siccome risulta $u_0(x_0) = 1$, si ha poi $u_0'(x_0) = \|x_0\|$. D'altra parte per ogni $x = \lambda x_0 \in H - \{0\}$ si ha:

$$\frac{|u_0'(x)|}{\|x\|} = \frac{|\lambda| \cdot \|x_0\|}{\|\lambda x_0\|} = 1 ,$$

sicché u_0' è continuo, e risulta:

$$\|u_0'\|_{H^*} = 1 .$$

Per il teorema di Hahn-Banach esiste allora un prolungamento u_1 di u_0' su S, lineare e continuo, tale che:

$$\|u_1\|_{S^*} = 1 .$$

Abbiamo con ciò stabilito che per ogni $x_0 \in S - \{0\}$ esiste un elemento u_1 di S^* avente norma 1 e soddisfacente alla (4) con $x = x_0$. Abbiamo così dimostrato la seconda affermazione dell'enunciato.

Se ne deduce subito la prima parte del teorema. Scelto infatti un numero positivo α e detto u_1 il funzionale che per il fissato $x \in S - \{0\}$ verifica la (4) ed è tale che $\|u_1\| = 1$, il funzionale $u = \alpha u_1$ verifica la (3) in quanto $\|u\| = \alpha \|u_1\| = \alpha$.

(9.4). Se x_1 ed x_2 sono punti distinti dello spazio vettoriale normato S, esiste almeno un funzionale lineare u, definito e continuo in S, tale che $u(x_1) \neq u(x_2)$.

Dim. Se $x_2 = 0$, dal teor. (9.2) [con $H = \{0\}$ e x_1 in luogo di x_0] consegue che esiste un funzionale lineare u, definito e continuo in S, tale che $u(x_1) = 1$. Essendo $u(x_2) = u(0) = 0$, ciò prova l'asserto nel caso in esame. Passando al caso generale, posto $y = x_1 - x_2 \neq 0$ esiste, in virtù di quanto già dimostrato, un funzionale lineare u, definito e continuo in S, tale che $u(y) \neq u(0)$, ossia $u(y) \neq 0$. Ciò vuol dire che $u(x_1) \neq u(x_2)$ e il teorema è dimostrato.

Utilizzando il corollario (9.2), dimostriamo il teorema seguente:

(9.5). Sia S uno spazio vettoriale normato.

Qualunque siano i punti di S linearmente indipendenti:

(7) $$x_1, x_2, \dots, x_n$$

esistono n funzionali definiti in S.

(8) $$u_1, u_2, \dots, u_n$$

lineari e continui, tali che:

(9) $$u_i(x_k) = \delta_{ik} \qquad \forall i,k \in \{1,\dots,n\}$$ (1)

Reciprocamente qualunque siano i funzionali lineari e continui (8) definiti in S, linearmente indipendenti, esistono n punti (7) di S tali da verificare la (9).

Dim. Assegnati i punti (7) linearmente indipendenti, per ogni $k \in \{1,\dots,n\}$ sia H_k il sottospazio di S generato da $\{x_1,\dots,x_n\} - \{x_k\}$: in virtù della (9.2), esiste almeno un funzionale u_k definito in S, lineare e continuo, tale che $u_k(x_k) = 1$ e identicamente nullo in H_k [quindi tale che $u_k(x_i) = 0$ per $i \neq k$]. I funzionali $u_1, \dots, u_n$ così ottenuti soddisfano alla (9).

Reciprocamente, siano assegnati i funzionali (8) lineari e continui in S, linearmente indipendenti.

Proviamo dapprima che esistono n punti di S:

$$y_1, \dots, y_n$$

tali che il determinante:

$$\Delta_n = \begin{vmatrix} u_1(y_1) & \dots & u_n(y_1) \\ \dots & \dots & \dots \\ u_1(y_n) & \dots & u_n(y_n) \end{vmatrix}$$

sia diverso di zero.

Tale asserzione è ovvia se $n=1$, giacché u_1 non è identicamente nullo; supposto che sia vera per $n=k$, proviamo che è vera altresì per $n=k+1$. Siano $y_1, \dots, y_k \in S$ tali che $\Delta_k \neq 0$: si tratta di far vedere che esiste $y_{k+1} \in S$ tale che $\Delta_{k+1} \neq 0$. Procedendo per assurdo, ammettiamo che $\Delta_{k+1} = 0 \ \forall y_{k+1} \in S$: sviluppando il determinante Δ_{k+1} secondo gli elementi dell'ultima riga, si avrebbe:

$$\sum_{i=1}^{k+1} \lambda_i u_i(y_{k+1}) = 0$$

$\lambda_1, \dots, \lambda_{k+1}$, essendo degli scalari, con $\lambda_{k+1} = \Delta_k \neq 0$. Conseguentemente $u_1, \dots, u_{k+1}$ sarebbero, contro il supposto, linearmente dipendenti.

Scelti $y_1, \dots, y_n$ tali che $\Delta_n \neq 0$, per ogni $k \in \{1,\dots,n\}$ esistono n scalari $\alpha_{k1}, \dots, \alpha_{kn}$ tali che:

$$\sum_{j=1}^{n} \alpha_{kj}\, u_i(y_j) = \delta_{ik} \qquad \forall i \in \{1,\dots,n\},$$

ossia tali che:

$$(10) \qquad u_i\left(\sum_{j=1}^{n} \alpha_{kj}\, y_j\right) = \delta_{ik} \qquad \forall i \in \{1,\dots,n\}.$$

Posto:

$$x_k = \sum_{j=1}^{n} \alpha_{kj}\, y_j \qquad \forall k \in \{1,\dots,n\},$$

in virtù della (10) i punti $x_1,\dots,x_n$ soddisfano alla (9).

— — — — — — — — — —

10. Biduale di uno spazio normato. Spazi riflessivi. Spazi di Banach uniformemente convessi.

Sia S uno spazio vettoriale normato.

In un primo tempo supponemo che lo spazio vettoriale sostegno di S sia sul campo reale, sicché la funzione di due variabili:

$$(1) \qquad \langle u, x \rangle \quad ,$$

definita per $u \in S^*$ ed $x \in S$, è una forma bilineare in $S^* \times S$ (che si chiama, come si è detto a pag. 172, la dualità tra S ed S^*).

Abbiamo anche detto che i simboli:

$$(2) \qquad \langle \cdot, x \rangle \quad , \quad \langle u, \cdot \rangle$$

rappresentano due funzionali lineari: il primo, nel quale x è un fissato elemento di S, denota il funzionale che ad ogni $u \in S^*$ associa il numero reale (1), e analogamente il secondo, nel quale u è un fissato elemento di S^*, denota il funzionale che ad ogni $x \in S$ associa il numero reale (1) [ossia, denota il funzionale u].

In virtù della disuguaglianza [cfr. (5') pag. 174]:

$$|\langle u, x \rangle| \leq \|u\| \cdot \|x\| \qquad \forall u \in S^* \text{ e } \forall x \in S,$$

per il teor. (2.1) di pag. 138 i funzionali lineari (2) sono entrambi continui. Come il secondo è un elemento dello spazio (vettoriale normato) S^*, duale di S, così il primo, essendo un funzionale lineare e continuo definito in S^*, è un elemento dello spazio duale di S^*, che si chiama <u>il biduale di S</u> e si denota con S^{**}.

Peraltro, mentre il secondo dei funzionali (2) al variare di u in S^* descrive tutto lo spazio S^* (circostanza banale), non è detto che il primo, al variare di x in S, descriva tutto lo spazio S^{**}: orbene, quando quest'ultima circostanza si verifica, si dice

che lo spazio vettoriale normato S è <u>riflessivo</u>.

Abbiamo supposto che il sostegno di S fosse uno spazio vettoriale sul campo reale solo per rendere più scorrevole il discorso, dato che nel caso S complesso gli elementi del duale (e quindi anche quelli del biduale) non sono funzionali lineari, ma sono <u>coniugati</u> di funzionali lineari. Nel caso S complesso il primo dei funzionali (2) rappresenta un funzionale lineare (e continuo), e non <u>il coniugato</u> di un funzionale lineare, per cui non è un elemento dello spazio $S^{\times\times}$ biduale di S. Un elemento di $S^{\times\times}$ è invece il suo coniugato, che naturalmente si denoterà con

$$\overline{\langle \cdot , x \rangle} \tag{3}$$

Con riferimento ad uno spazio vettoriale normato S reale o complesso, l'operatore J che ad ogni $x \in S$ associa l'elemento (3) di $S^{\times\times}$:

$$J : x \in S \to \overline{\langle \cdot , x \rangle} \in S^{\times\times} \tag{4}$$

si chiama <u>l'applicazione canonica di S in $S^{\times\times}$</u>. Lo spazio S si dice dunque <u>riflessivo</u> quando tale applicazione è suriettiva, cioè quando ogni funzionale $v \in S^{\times\times}$ è del tipo (3) con $x \in S$.

Ricordando la definizione della norma nel duale topologico di uno spazio vettoriale normato [pag. 173], per ogni $v \in S^{\times\times}$ si ha per definizione:

$$\|v\|_{S^{\times\times}} = \sup_{u \in S^{\times} - \{0\}} \frac{|\langle\langle v, u \rangle\rangle|}{\|u\|}$$

dove abbiamo indicato con $\langle\langle v,u \rangle\rangle$ la dualità tra $S^{\times}$ ed $S^{\times\times}$ (sicché $\langle\langle v,u \rangle\rangle$ significa $v(u)$), anche se, solitamente, si denota anch'essa con $\langle v,u \rangle$. Fissato $x \in S$ e posto $v = Jx$, si ha:

$$\langle\langle Jx, u \rangle\rangle = \overline{\langle u, x \rangle} \quad \forall u \in S^{\times}, \tag{5}$$

e quindi:

$$\|Jx\|_{S^{\times\times}} = \sup_{u \in S^{\times} - \{0\}} \frac{|\langle\langle Jx, u \rangle\rangle|}{\|u\|} = \sup_{u \in S^{\times} - \{0\}} \frac{|\langle u, x \rangle|}{\|u\|}$$

da cui, ricordando la (6) di pag. 184, consegue:

$$\|Jx\|_{S^{\times\times}} = \|x\|_S \qquad \forall x \in S . \tag{6}$$

Pertanto l'applicazione canonica di S in $S^{\times\times}$, ovviamente lineare, è un isomorfismo isometrico di S sul sottospazio $J(S)$ di $S^{\times\times}$. Se S è riflessivo, tale applicazione si chiama <u>l'isomorfismo canonico di S su $S^{\times\times}$</u>.

Dunque se S è riflessivo lo spazio $S^{\times\times}$ biduale di S è equivalente ad S, e quindi si può identificare con S.

Avvertiamo che è possibile costruire esempi di spazi S non riflessivi per i quali esiste un isomorfismo isometrico di S su S^{**}. Può accadere cioè che l'operatore (4) non sia suriettivo e che esista un altro operatore di S in S^{**} suriettivo, anch'esso isomorfismo isometrico: in tal caso lo spazio S non si dice riflessivo. Pertanto dire che S è riflessivo non equivale a dire che lo spazio S è equivalente al suo biduale, ma significa "qualcosa in più": precisamente, nel caso S reale, se si identificano gli spazi S ed S^{**} mediante l'isomorfismo canonico, la forma bilineare $\langle \cdot, \cdot \rangle$ rappresenta sia la dualità tra S ed S^* che la dualità tra S^* ed S^{**} [pensando $\langle u, x \rangle$ anche come il valore "del funzionale x" nel punto u]; nel caso S complesso, la forma sesquilineare $\langle \cdot, \cdot \rangle$ rappresenta la dualità tra S ed S^*, e la sua coniugata la dualità tra S^* ed S^{**}.

— — — — — — — — — — —

Consideriamo ad esempio due numeri reali p, p', maggiori di 1 e tali che $\frac{1}{p} + \frac{1}{p'} = 1$. Come si è visto nell'es. 4 di pag. 181, il duale dello spazio $L^p(\Omega)$ si può identificare con $L^{p'}(\Omega)$, e per lo stesso motivo il duale di $L^{p'}(\Omega)$ si può identificare con $L^p(\Omega)$. Con tali identificazioni, se $x \in L^p(\Omega)$ ed $y \in L^{p'}(\Omega)$ l'integrale:

$$(7) \qquad \int_\Omega x(t) y(t)\, dt \ ,$$

che è una forma bilineare in $L^p(\Omega) \times L^{p'}(\Omega)$, rappresenta sia la dualità tra $S = L^p(\Omega)$ ed $S^* = L^{p'}(\Omega)$, sia la dualità tra $S^* = L^{p'}(\Omega)$ ed $S^{**} = L^p(\Omega)$. In altri termini: fissato $x \in L^p(\Omega)$, l'integrale (7) come funzione di y è un elemento del duale di $L^{p'}(\Omega)$, e al variare di x fornisce tutti gli elementi del predetto duale; fissato $y \in L^{p'}(\Omega)$, l'integrale (7) come funzione di x è un elemento del duale di $L^p(\Omega)$, e al variare di y fornisce tutti gli elementi del duale di $L^p(\Omega)$.

Perciò lo spazio $S = L^p(\Omega)$, che è equivalente al suo biduale, è uno spazio riflessivo.

Prescindendo dalle suddette identificazioni, per ogni $v \in S^{**} = (L^p(\Omega))^{**}$ esiste uno ed un solo $x \in L^p(\Omega)$ tale che $Jx = v$, e quindi tale che:

$$\langle\langle v, u \rangle\rangle = \langle u, x \rangle \quad \forall u \in S^*,$$

e d'altra parte per ogni $u \in S^* = (L^p(\Omega))^*$ esiste uno ed un solo $y \in L^{p'}(\Omega)$ tale che:

$$(8) \qquad \langle u, x \rangle = \int_\Omega x(t) y(t)\, dt \ .$$

Di conseguenza risulta:

$$(9) \qquad \langle\langle v, u \rangle\rangle = \int_\Omega x(t) y(t)\, dt \ ,$$

sicchè la forma bilineare (7) rappresenta sia la dualità tra S ed S^* [cfr. (8)], sia la dualità tra S^* ed S^{**} [cfr. (9)].

Notiamo che, posto $y = Au$, è preferibile scrivere le (8) e (9) nella forma:

$$\langle A\ddot{y}, x \rangle = \int_\Omega x(t) y(t)\, dt \ , \qquad \langle\langle Jx, A\ddot{y} \rangle\rangle = \int_\Omega x(t) y(t)\, dt$$

in modo che a primo membro figurino le variabili x, y che figurano a secondo membro.

— — — — — — — — — — —

Quando lo spazio S non è riflessivo, si può ugualmente identificare ciascun elemento x di S con l'elemento Jx del biduale S^{**}, e quindi pensare

$\langle u, x \rangle$ come il valore "del funzionale x" nel punto u, però la forma sesquilineare $\langle u, x \rangle$ non è allora definita in $S^* \times S^{**}$, ma in $S^* \times J(S)$ [e quindi non rappresenta la dualità tra S^* ed S^{**}].

Poichè S^{**} è il duale di S^*, dal teor. (6.2) di pag. 174 consegue che S^{**} è in ogni caso uno spazio di Banach, sicchè uno spazio vettoriale normato S non completo non può essere riflessivo [altrimenti uno spazio normato non completo S sarebbe equivalente ad uno spazio completo, ossia ad S^{**}]. E' questo il motivo per cui si preferisce dire spazio di Banach riflessivo invece di "spazio vettoriale normato riflessivo".

Osserviamo che:

(10.1). Ogni spazio di Hilbert è riflessivo.

Dim. Siccome S^* è uno spazio di Hilbert [cfr. (7.4) pag. 180], per il teorema di Riesz [cfr. (7.1) pag. 177] applicato con S^* in luogo di S, per ogni $v \in S^{**}$ esiste uno ed un solo elemento w di S^* tale che:

$$v(u) = (w, u)_{S^*} \quad \forall u \in S^* . \tag{10}$$

Orbene, detto j l'isomorfismo canonico di S^* in S e posto $jw = x$, si verifica subito che:

$$v = Jx = \overline{\langle \cdot, x \rangle},$$

cioè che:

$$v(u) = \overline{\langle u, x \rangle} \quad \forall u \in S^* .$$

Infatti dalla (10) consegue, qualunque sia $u \in S^*$:

$$v(u) = (w, u)_{S^*} = (jw, ju)_S = \overline{(ju, x)} = \overline{\langle u, x \rangle} \qquad (1).$$

Pertanto l'operatore (4) è suriettivo, e ciò vuol dire che lo spazio di Hilbert S è riflessivo.

Riportiamo senza dimostrazione i teoremi seguenti:

(10.2): Se S è uno spazio di Banach riflessivo, ogni spazio di Banach equivalente ad S è riflessivo.

(10.3). Qualunque sia lo spazio di Banach S, sussiste l'equivalenza:

$$S \text{ riflessivo} \iff S^* \text{ riflessivo}.$$

(10.4). Sia S uno spazio di Banach, e sia H un sottospazio vettoriale di S, chiuso in S. Se S è riflessivo, lo spazio di Banach di sostegno H è riflessivo.

---

(1) Ricordiamo [cfr. (9) pag. 180] che nel duale S^* di uno spazio di Hilbert S il prodotto scalare è definito da:

$$(u_1, u_2)_{S^*} = (ju_1, ju_2)_S .$$

(10.5). Ogni spazio vettoriale normato a dimensione finita e' riflessivo.

(10.6). Qualunque siano $p \in]1,+\infty[$, $m \in N$, gli spazi normati:

$$l_p \ , \quad L^p(\Omega) \ , \quad W^{m,p}(\Omega) \ , \quad H^{m,p}(\Omega) \ , \quad H_0^{m,p}(\Omega)$$

sono riflessivi.

Gli spazi l_1 , l_∞ , $L^1(\Omega)$, $L^\infty(\Omega)$ non sono riflessivi.

(10.7). Qualunque siano $m \in N$ e $\lambda \in]0,1]$, gli spazi normati:

$$C^m(\bar{\Omega}) \ , \quad C^{m,\lambda}(\bar{\Omega})$$

non sono riflessivi.

Vogliamo infine segnalare una condizione sufficiente affinche' uno spazio di Banach sia riflessivo. A tale scopo osserviamo preliminarmente quanto segue.

Se S e' uno spazio di Hilbert, dette $(e'_n)_{n\in N}$, $(e''_n)_{n\in N}$ due successioni di punti di S aventi norma uguale ad 1, per la (4) di pag. 97 si ha:

$$\| e'_n + e''_n \|^2 + \| e'_n - e''_n \|^2 = 4 \qquad \forall n \in N$$

e quindi per tali successioni sussiste l'implicazione:

$$(11) \qquad \lim \| e'_n - e''_n \| = 2 \implies \lim \| e'_n + e''_n \| = 0 \ .$$

Se la norma in uno spazio di Banach S non e' generata da un prodotto scalare [nel qual caso la citata (4) di pag. 97 non sussiste], non e' detto che sia vera la (11), sempre con riferimento a successioni di punti di norma unitaria. Perciò quando sussiste la (11) lo spazio S gode di una proprietà particolare, che si chiama "uniforme convessità".

Dunque uno spazio di Banach S si dice uniformemente convesso quando, con riferimento a successioni $(e'_n)_{n\in N}$, $(e''_n)_{n\in N}$ di punti di S aventi norma uguale ad 1, sussiste l'implicazione (11).

Ad esempio, si dimostra che qualunque siano $p \in]1,+\infty[$, $m \in N$, gli spazi:

$$l_p \ , \quad L^p(\Omega) \ , \quad W^{m,p}(\Omega) \ , \quad H^{m,p}(\Omega) \ , \quad H_0^{m,p}(\Omega)$$

sono uniformemente convessi.

Orbene, sussiste il seguente teorema, di cui omettiamo la dimostrazione:

(10.8). Ogni spazio di Banach uniformemente convesso e' riflessivo.

11. Topologia della convergenza puntuale nel duale di uno spazio normato.

Sia S uno spazio vettoriale normato.

Come abbiamo visto, al duale topologico di S si associa una norma, e si ottiene così il duale forte di S che brevemente si chiama il duale di S.

Al duale topologico di S si associa anche un'altra topologia, oltre a quella generata dalla norma: la topologia della convergenza puntuale.

Si tratta della topologia costruita come quella del n. 6 di pag. 30: basta sostituire ivi l'insieme S con S^* e il sottoinsieme A di R^X con il sostegno di S. Siccome nella costruzione di tale topologia, topologia per S^*, intervengono gli elementi di S, noi la denoteremo col simbolo:

$$\tau(S^*, S).$$

Da quanto detto a pag. 30 si deduce che, se $(u_n)_{n\in N}$ è una successione di punti di S^* e se u_0 è un punto di S^*, la relazione:

(1) $\lim u_n = u_0$ nello spazio topologico $(S^*, \tau(S^*,S))$

equivale al fatto che risulta:

(2) $$\lim_{n\to\infty} u_n(x) = u_0(x) \quad \forall x \in S$$

ossia equivale al fatto che la $(u_n)_{n\in N}$, come successione di funzioni definite in S, converge puntualmente in S alla funzione u_0. Perciò la topologia $\tau(S^*,S)$ si chiama la topologia della convergenza puntuale, e la (1) si esprime dicendo che la successione $(u_n)_{n\in N}$ di elementi di S^* converge puntualmente ad $u_0 \in S^*$; l'elemento u_0 di S^* si chiama, di conseguenza, il limite puntuale della successione.

Se per ogni $x \in S$ esiste finito il limite a primo membro della (2), si dice ancora che la successione $(u_n)_{n\in N}$ converge puntualmente. In tal caso si può considerare il funzionale:

(3) $$u_0 : x \in S \longrightarrow \lim_{n\to\infty} u_n(x),$$

sicché la (2) sussiste per definizione di u_0; però non è detto che il funzionale u_0 [che si chiama ancora il limite puntuale della successione] sia un elemento di S^*, perché u_0 è ovviamente un funzionale lineare se S è reale, è il coniugato di un funzionale lineare se S è complesso, ma non è detto che sia continuo.

Dunque se una successione $(u_n)_{n\in N}$ di elementi di S^* è puntualmente convergente, non è detto che sussista la (1). In proposito osserviamo che:

(11.1). Se la successione $(u_n)_{n\in N}$ di elementi di S^* è puntualmente convergente ed è limitata, il suo limite puntuale u_0 appartiene ad S^*.

Dim. Si tratta di dimostrare la continuità del funzionale (3). Per la (5') di pag. 176 si ha:

$$|\langle u_n, x\rangle| \le \|u_n\|\cdot\|x\| \quad \forall n\in N \text{ e } \forall x\in S$$

e quindi, se $\|u_n\|\le k$ $\forall n\in N$, risulta:

$$|\langle u_n, x\rangle| \le k\|x\| \quad \forall n\in N \text{ e } \forall x\in S.$$

Passando al limite per $n\to\infty$ ne segue che:

$$|u_0(x)| \le k\|x\| \quad \forall x\in S$$

e pertanto il funzionale u_0 è continuo.

L'ipotesi di limitatezza che figura nella (11.1) è verificata automaticamente se lo spazio vettoriale normato S è completo. Sussiste inverso il teorema:

(11.2). Se S è uno spazio di Banach, ogni successione di elementi di S^* puntualmente convergente è limitata.

Dim. Nell'ipotesi che S sia uno spazio di Banach, consideriamo una successione $(u_n)_{n\in N}$ di elementi di S^*, puntualmente convergente in S.

Allo scopo di far vedere che esiste $l>0$ tale che:

$$\|u_n\|\le l \quad \forall n\in N, \tag{4}$$

dimostriamo in primo luogo che:

(5) Esistono $h>0$, $r>0$ e $x_0\in S$ tali che:
$$|\langle u_n, x\rangle| \le h \quad \forall n\in N \text{ e } \forall x\in C(x_0, r).$$

Procedendo per assurdo, neghiamo tale affermazione. Qualunque siano $x_0\in S$, $r_0>0$, esistono allora $x_1\in C(x_0,r_0)$ e $n_1\in N$ tali che $|\langle u_{n_1}, x_1\rangle|>1$; per la continuità di u_{n_1}, ciò implica che esiste $r_1>0$ tale che:

$$|\langle u_{n_1}, x\rangle| > 1 \quad \forall x\in C[x_1, r_1].$$

Analogamente, esistono $x_2\in C(x_1,r_1)$ e $n_2\in N$ tali che $|\langle u_{n_2}, x_2\rangle|>2$; per la continuità di u_{n_2}, ciò implica che esiste $r_2>0$ tale che:

$$|\langle u_{n_2}, x\rangle| > 2 \quad \forall x\in C[x_2, r_2].$$

Supponiamo di avere scelto r_i $(i=1,2)$ in modo che:

$$r_i < r_{i-1}/2 \quad , \qquad r_i < r_{i-1} - \|x_i - x_{i-1}\|.$$

Procedendo allo stesso modo, si definiscono induttivamente una successione $(n_k)_{k\in N}$ di interi positivi, una successione $(x_k)_{k\in N}$ di punti di S e una successione $(r_k)_{k\in N}$ di numeri positivi, tali da risultare, per ogni $k\in N$:

$$(6) \qquad |\langle u_{n_k}, x\rangle| > k \qquad \forall x \in C[x_k, r_k]$$

$$(7) \qquad r_k < r_{k-1}/2 \quad , \quad r_k < r_{k-1} - \|x_k - x_{k-1}\| .$$

Tenendo conto delle (7), si ha, qualunque siano $k, p \in N$:

$$(8) \qquad \|x_{k+p} - x_k\| \le \sum_{i=1}^{p} \|x_{k+i} - x_{k+i-1}\| < \sum_{i=1}^{p} (r_{k+i-1} - r_{k+i}) =$$
$$= r_k - r_{k+p} < r_k < r_0/2^k$$

e quindi $(x_k)_{k\in N}$ é una successione di Cauchy di S . Poiché S é completo, esiste il limite di $(x_k)_{k\in N}$, diciamolo $\bar{x}$. Dalla (8) consegue, passando al limite per $p \to \infty$:

$$\|\bar{x} - x_k\| \le r_k \qquad \forall k \in N$$

ossia $\bar{x} \in C[x_k, r_k]$ $\forall k \in N$. Per la (6) si ha allora:

$$|\langle u_{n_k}, \bar{x}\rangle| > k \qquad \forall k \in N$$

e quindi la successione numerica:

$$\left(|\langle u_n, \bar{x}\rangle|\right)_{n\in N}$$

non é limitata. Ma ciò é assurdo, in quanto tale successione é per ipotesi convergente.

Pertanto abbiamo acquisito la (5) . Mostriamo che dalla (5) segue la (4).

Qualunque sia $y \in S - \{0\}$, l'elemento di S :

$$x = x_0 + \frac{r}{2\|y\|} y$$

appartiene a $C(x_0, r)$, e quindi per la (5) risulta, per ogni $n \in N$:

$$\left|\langle u_n, x_0 + \frac{r}{2\|y\|} y\rangle\right| \le h$$

ossia:

$$\left|\frac{r}{2\|y\|}\langle u_n, y\rangle + \langle u_n, x_0\rangle\right| \le h$$

nonché, a maggior ragione:

$$\frac{r}{2\|y\|}\,|\langle u_n, y\rangle| - |\langle u_n, x_0\rangle| \le h .$$

Da tale relazione consegue, tenendo presente che $|\langle u_n, x_0\rangle| \le h$ [per la (5)]:

$$|\langle u_n, y\rangle| \le \frac{4h}{r}\|y\| \qquad \forall n \in N$$

e quindi, per la definizione di $\|u_n\|$, vale la (4) con $l = 4h/r$.

— — — — — — — — — — —

Un importante complemento al teorema ora dimostrato e' il seguente:

(11.3). Sia S uno spazio vettoriale normato. Se S e' separabile, da ogni successione limitata di punti di S^* si può estrarre una successione puntualmente convergente (ad un elemento di S^*).

Dim. Sia $(u_n)_{n\in N}$ una successione limitata di elementi di S^*, sicché esiste $l>0$ tale da risultare $\|u_n\| \le l$ $\forall n\in N$, ossia:

$$|\langle u_n, x\rangle| \le l\|x\| \quad \forall x\in S \text{ e } \forall n\in N. \tag{9}$$

Denotiamo con $(x_k)_{k\in N}$ una successione di punti di S il cui insieme di elementi sia ovunque denso in S.

In virtù della (9), per ogni $n\in N$ la successione numerica $(u_n(x_k))_{n\in N}$ è limitata, e quindi si può estrarre da essa (nonché da ogni sua estratta) una successione convergente. Sia dunque $(u_{r_1(n)})_{n\in N}$ una successione estratta da $(u_n)_{n\in N}$ convergente nel punto x_1; sia inoltre $(u_{r_2(n)})_{n\in N}$ una successione estratta da $(u_{r_1(n)})_{n\in N}$ convergente nel punto x_2, e quindi anche nel punto x_1. Procedendo analogamente, per ogni $k\in N$ si costruisce induttivamente una successione $(u_{r_k(n)})_{n\in N}$, estratta da $(u_{r_{k-1}(n)})_{n\in N}$, convergente nei punti $x_1, x_2, \dots, x_k$.

Si verifica facilmente che la successione "diagonale":

$$(u_{r_n(n)})_{n\in N} \tag{10}$$

è estratta dalla $(u_n)_{n\in N}$ e converge nel punto x_k qualunque sia $k\in N$.

Proviamo che, qualunque sia $x\in S$, la (10) converge nel punto x. Per semplicità di scrittura, poniamo $r_n(n) = h_n$ $\forall n\in N$.

Qualunque sia $\varepsilon>0$, poiché l'insieme degli elementi di $(x_k)_{k\in N}$ è ovunque denso in S, esiste $k\in N$ tale che:

$$\|x - x_k\| < \varepsilon/3l ;$$

determinato k, poiché $(u_{h_n})_{n\in N}$ converge nel punto x_k, esiste $\nu\in N$ tale che:

$$|u_{h_p}(x_k) - u_{h_q}(x_k)| < \varepsilon/3 \quad \forall p>\nu \text{ e } \forall q>\nu .$$

Ne segue, tenendo conto anche della (9), che qualunque siano $p,q>\nu$ risulta:

$$|u_{h_p}(x) - u_{h_q}(x)| \le |u_{h_p}(x) - u_{h_p}(x_k)| + |u_{h_p}(x_k) - u_{h_q}(x_k)| + |u_{h_q}(x_k) - u_{h_q}(x)| \le$$
$$\le l\|x - x_k\| + \frac{\varepsilon}{3} + l\|x - x_k\| < \varepsilon .$$

Pertanto la successione $(u_{h_n}(x))_{n\in N}$ è una successione (numerica) di Cauchy, e quindi è convergente.

In definitiva la successione $(u_{h_n})_{n\in N}$, estratta dalla $(u_n)_{n\in N}$, è puntualmente convergente in S.

Avvertiamo che i vari Autori non adoperano tutti la stessa denominazione per la topologia della convergenza puntuale $\tau(S^*, S)$: alcuni la chiamano <u>S-topologia per S^*</u>, altri topologia <u>* debole</u>, altri ancora <u>topologia indebolita</u>.

Si dimostra che, associando al duale topologico di S tale topologia, si ottiene uno spazio vettoriale topologico di Hausdorff localmente convesso.

12. Topologia debole per uno spazio vettoriale normato.

Sia S uno spazio vettoriale normato.

Al sostegno di S si suole associare un'altra topologia (oltre a quella generata dalla norma) che si chiama la topologia debole per S, e che si costruisce come segue.

Detto x_0 un punto di S, scegliamo un numero positivo ε, un intero positivo k, nonché k elementi di S^*, diciamoli $u_1, \dots, u_k$. Consideriamo quindi le parti di S che contengono l'insieme:

$$(1) \qquad \{x \in S : |u_i(x) - u_i(x_0)| < \varepsilon \ \ \forall i \in \{1, \dots, k\}\}$$

Al variare di ε in $]0, +\infty[$, di k in N, e di $u_1, \dots, u_k$ in S^*, le predette parti di S descrivono un insieme di parti di S, che è individuato dal punto x_0: diciamolo $\mathcal{J}(x_0)$. Orbene, si verifica facilmente che la famiglia $(\mathcal{J}(x))_{x \in S}$ soddisfa agli assiomi di Hausdorff, e quindi esiste una e una sola topologia per il sostegno di S tale che, per ogni $x \in S$, $\mathcal{J}(x)$ è l'insieme degli intorni di x.

Nella costruzione di tale topologia intervengono gli elementi di S^*, e quindi indirettamente vi interviene anche la norma dello spazio S al cui sostegno essa si riferisce; perciò noi la indicheremo col simbolo:

$$(2) \qquad \mathcal{T}(S, S^*).$$

Si dimostra che, associando al sostegno di S la topologia (2), si ottiene uno spazio vettoriale topologico di Hausdorff localmente convesso, che noi denoteremo con S_d.

Osserviamo che:

(12.1) La topologia $\mathcal{T}(S, S^*)$ soddisfa alle seguenti condizioni:

1) È più debole della topologia generata dalla norma.

2) Rispetto ad essa tutti gli elementi di S^* sono continui [1].

3) È la più debole fra tutte le topologie rispetto alle quali gli elementi di S^* sono continui.

Dim. 1). Detto x_0 un punto di S, sia I un intorno di x_0 nello spazio S_d, e siano $\varepsilon, k, u_1, \dots, u_k$ tali che I contenga l'insieme (1). Detto h un numero maggiore delle norme in S^* di $u_1, u_2, \dots, u_k$, sia x un arbitrario punto del cerchio aperto di S di centro x_0 e raggio ε/h, sicché risulta $\|x - x_0\| < \varepsilon/h$. Si ha:

---

[1] Cioè per ogni $u \in S^*$ il funzionale $u : S_d \to K$ è continuo. Come al solito, K denota l'insieme dei numeri reali o quello dei numeri complessi (considerato come spazio vettoriale normato).

$$|u_i(x)-u_i(x_0)| = |u_i(x-x_0)| \le \|u_i\|\cdot\|x-x_0\| \le h\|x-x_0\| < \varepsilon \qquad \forall i \in \{1,\dots,\kappa\},$$

e quindi x appartiene all'insieme (1), nonché di conseguenza all'intorno I di x_0 in S_d. Dunque I contiene un cerchio aperto di centro x_0 dello spazio S, e pertanto I è anche un intorno di x_0 in S. Ne segue che ogni aperto di S_d è anche aperto nello spazio S, e ciò prova la 1).

Dim.2). Sia $u \in S^*$ e sia $x_0 \in S$. Per ogni $\varepsilon > 0$ l'intorno di x_0 in S_d:

$$I = \{ x \in S : |u(x)-u(x_0)| < \varepsilon \}$$

è tale che:

$$x \in I \implies |u(x)-u(x_0)| < \varepsilon .$$

Pertanto il funzionale $u : S_d \to K$ è continuo in x_0.

Dim.3). Sia $\mathcal{T}$ una topologia rispetto alla quale gli elementi di S^* sono continui, e sia S_1 lo spazio topologico che si ottiene associando al sostegno di S la topologia $\mathcal{T}$.

Detto x_0 un punto di S, sia I un intorno di x_0 in S_d, e siano $\varepsilon, \kappa, u_1, \dots, u_\kappa$ tali che I contenga l'insieme (1). Essendo $u_1 \dots u_\kappa$ continui rispetto a $\mathcal{T}$, esiste un intorno I_1 di x_0 in S_1 tale che:

$$x \in I_1 \implies |u_i(x)-u_i(x_0)| < \varepsilon \quad \forall i \in \{1,\dots,\kappa\},$$

sicché I_1 è incluso nell'insieme (1), quindi anche in I. Siccome I contiene un intorno di x_0 in S_1, l'insieme I è esso stesso un intorno di x_0 in S_1.

Dunque ogni intorno I di x_0 in S_d è anche intorno di x_0 in S_1, e da ciò segue che la topologia (2) è più debole della topologia $\mathcal{T}$.

L'osservazione (12.1) giustifica il fatto che la topologia $\mathcal{T}(S,S^*)$ si chiama <u>la topologia debole per S</u>. In contrapposto la topologia generata dalla norma (topologia "iniziale") viene detta <u>la topologia forte per S</u>. Si dimostra che:

(12.2). <u>Se lo spazio vettoriale normato S è a dimensione finita, la topologia debole e la topologia forte per S coincidono</u>.

Osserviamo ora che:

(12.3). <u>Per ogni successione $(x_n)_{n\in\mathbb{N}}$ di punti di S e per ogni $x_0 \in S$, la relazione:</u>

(3) $$\lim x_n = x_0 \quad \text{in } S_d$$

<u>equivale alla seguente</u>:

(4) $$\lim \langle u, x_n \rangle = \langle u, x_0 \rangle \quad \forall u \in S^* .$$

Dim. Nell'ipotesi (3), sia u un arbitrario elemento di S^*. Per ogni $\varepsilon > 0$, detto I l'intorno di x_0 in S_d:

$$I = \{ x \in S : |u(x)-u(x_0)| < \varepsilon \},$$

per la (3) esiste $\nu \in \mathbb{N}$ tale che:

$$x_n \in I \quad \forall n > \nu .$$

Dunque per ogni $n > \nu$ risulta:

$$|u(x_n)-u(x_0)| < \varepsilon$$

e ciò prova la (4).

Viceversa, supponiamo che sussista la (4) e denotiamo con I un intorno di x_0 in S_d : siano $\varepsilon, k, u_1, \dots, u_k$ tali che I contenga l'insieme (1). Essendo per la (4):

$$\lim u_i(x_n) = u_i(x_0) \quad \forall i \in \{1, \dots, k\},$$

esiste $\nu \in N$ tale che:

$$|u_i(x_n) - u_i(x_0)| < \varepsilon \quad \forall n > \nu \text{ e } \forall i \in \{1, \dots, k\}.$$

Ciò significa che per ogni $n > \nu$ il punto x_n appartiene all'insieme (1), quindi anche ad I, e la (5) è così dimostrata.

— — — — — — — — — —

In virtù della (12.3.), quando si verifica la (4) si dice che la successione $(x_n)_{n \in N}$ converge debolmente ad x_0, e il punto x_0 si chiama il limite debole della successione. In contrapposto, se la successione converge ad x_0 nello spazio normato S, si dice che essa converge fortemente ad x_0.

Se per ogni $u \in S^*$ esiste finito il limite a primo membro della (4), si dice ancora che la successione $(x_n)_{n \in N}$ è debolmente convergente, ma ciò non garantisce l'esistenza di un punto x_0 di S tale che sussista la (4). In altri termini, se una successione è debolmente convergente in S, non è detto che essa sia dotata di limite debole. In proposito sussiste il teorema:

(12.4). Se S è uno spazio di Banach riflessivo, ogni successione di punti di S debolmente convergente è dotata di limite debole.

— — — — — — — — —

Dim. Sia $(x_n)_{n \in N}$ una successione di punti di S debolmente convergente. Detto J l'isomorfismo canonico di S su S^{**} [cfr. (4) pag. 187], qualunque sia $u \in S^*$ risulta:

$$(5) \qquad \langle\langle Jx_n, u \rangle\rangle = \overline{\langle u, x_n \rangle} \quad \forall n \in N$$

dove $\langle\langle \cdot, \cdot \rangle\rangle$ denota la dualità tra S^* ed S^{**}. Essendo $(x_n)_{n \in N}$ debolmente convergente, ne segue che la successione di elementi di S^{**}:

$$(6) \qquad (Jx_n)_{n \in N}$$

è puntualmente convergente in S^*. Poiché S^* è uno spazio di Banach [cfr. (6.2) pag. 174], per il teor. (11.2) di pag. 192 (applicato con S^* in luogo di S) la successione (6) è limitata, e quindi, per il teor. (11.1) di pag. 192, il suo limite puntuale, diciamolo v, appartiene ad S^{**}.

Essendo S riflessivo, esiste $x_0 \in S$ tale che $v = Jx_0$, ossia tale che:

$$\langle\langle v, u \rangle\rangle = \overline{\langle u, x_0 \rangle} \quad \forall u \in S^*,$$

e d'altra parte, per definizione di limite puntuale, si ha:

$$\langle\langle v, u \rangle\rangle = \lim \langle\langle Jx_n, u \rangle\rangle \quad \forall u \in S^*.$$

Pertanto risulta:

$$\lim \langle\langle Jx_n, u \rangle\rangle = \overline{\langle u, x_0 \rangle} \quad \forall u \in S^*$$

e quindi, per la (5):

$$\lim \langle u, x_n \rangle = \langle u, x_0 \rangle \quad \forall u \in S^*.$$

Dunque x_0 è il limite debole della successione $(x_n)_{n \in N}$, e l'assunto è dimostrato.

— — — — — — — — — —

Si verifica subito che:

(12.5). Qualunque sia lo spazio vettoriale normato S, ogni successione di punti di S debolmente convergente è limitata.

Dim. Detta J l'applicazione canonica di S in S^{**} [cfr. pag. 187], col ragionamento esposto nella prima parte del teorema precedente si conclude che la successione (x_n) è limitata, ossia che esiste $l \geq 0$ tale che $\|Jx_n\|_{S^{**}} \leq l \quad \forall n \in N$. Essendo [cfr. (6) pag. 187]:

$$\|Jx_n\|_{S^{**}} = \|x_n\|_S \quad \forall n \in N,$$

ne segue l'asserto.

Un importante complemento al teor. (12.5) è il teorema seguente, del quale omettiamo la dimostrazione:

(12.6). Se S è uno spazio di Banach riflessivo, da ogni successione limitata di punti di S se ne può estrarre una debolmente convergente (ad un punto di S (1)).

Ovviamente, per la (12.1) di pag. 195, proposiz. 1), una successione di punti di uno spazio normato S convergente fortemente ad x_0 converge anche debolmente ad x_0. Se S è uno spazio di Banach uniformemente convesso [cfr. pag. 190], si può precisare quale condizione occorre aggiungere alla convergenza debole per ottenere la convergenza forte. Sussiste invero il teorema:

(12.7). Sia S uno spazio di Banach uniformemente convesso. Se $(x_n)_{n \in N}$ è una successione di punti di S, e se x_0 è un punto di S, vale l'implicazione:

$$(7) \qquad \left(\begin{array}{c} \lim x_n = x_0 \ \text{debolmente} \\ \lim \|x_n\| = \|x_0\| \end{array} \right) \Longrightarrow \lim x_n = x_0 \ \text{fortemente}.$$

E' opportuno ricordare che la condizione $\|x_n\| \to \|x_0\|$ è necessaria affinchè la successione $(x_n)_{n \in N}$ converga fortemente ad x_0 [cfr. nota (1) pag. 163].

Anche del teor. (12.7) non riportiamo la dimostrazione. Vogliamo peraltro rilevare che il teorema è pressocchè evidente nel caso che S sia uno spazio di Hilbert (2).

---

(1) Cfr. (12.4).

(2) Dalla definizione introdotta a pag. 190 consegue banalmente che gli spazi di Hilbert sono uniformemente convessi.

Osserviamo dapprima che, se S è uno spazio di Hilbert e se $(x_n)_{n\in N}$ è una successione di punti di S, detto j l'isomorfismo canonico di S^* su S si ha [cfr. (7) pag. 179]:

$$\langle u, x_n\rangle = (ju, x_n) \quad \forall u \in S^*$$

e quindi la convergenza debole di $(x_n)_{n\in N}$ ad x_0 equivale al fatto che:

$$\lim_{n\to\infty} (x, x_n) = (x, x_0) \qquad \forall x \in S$$

conformemente a quanto si è detto a pag. 107.

Allora l'implicazione (7) consegue banalmente dall'uguaglianza:

$$\|x_n - x_0\|^2 = (x_n - x_0, x_n - x_0) = \|x_n\|^2 - (x_0, x_n) - (x_n, x_0) + \|x_0\|^2 \qquad \forall n \in N .$$

Osserviamo pure che:

(12.8). Se S è uno spazio di Hilbert, qualunque siano $x_0, y_0 \in S$ e $(x_n)_{n\in N}$, $(y_n)_{n\in N}$ successioni di punti di S, vale l'implicazione:

$$\left.\begin{array}{l} \lim x_n = x_0 \ \text{debolmente} \\ \lim y_n = y_0 \ \text{fortemente} \end{array}\right) \Longrightarrow \lim (x_n, y_n) = (x_0, y_0) .$$

- - - - - - - - - -

Dim. Se la successione $(x_n)_{n\in N}$ converge debolmente ad x_0, esiste $l > 0$ tale che $\|x_n\| \le l \ \forall n \in N$ [Teor. (12.5) pag. 198], ed inoltre risulta $\lim (x_n, y_0) = (x_0, y_0)$. L'assunto consegue allora dalla relazione:

$$|(x_n, y_n) - (x_0, y_0)| = |(x_n, y_n) - (x_n, y_0) + (x_n, y_0) - (x_0, y_0)| \le$$

$$\le |(x_n, y_n - y_0)| + |(x_n, y_0) - (x_0, y_0)| \le \|x_n\| \cdot \|y_n - y_0\| + |(x_n, y_0) - (x_0, y_0)| \le$$

$$\le l\, \|y_n - y_0\| + |(x_n, y_0) - (x_0, y_0)| .$$

- - - - - - - - - -

Riportiamo inoltre senza dimostrazione il seguente teorema, in base al quale un operatore compatto trasforma successioni debolmente convergenti in successioni fortemente convergenti:

(12.9). Siano S, S_1 spazi vettoriali normati, ed A un operatore lineare di S in S_1. Se A è compatto, per ogni successione $(x_n)_{n\in N}$ di punti di S e per ogni $x_0 \in S$ vale l'implicazione:

(8) $$\lim x_n = x_0 \ \text{debolmente} \Longrightarrow \lim Ax_n = Ax_0 \ \text{fortemente} .$$

Nell'ipotesi che S sia uno spazio di Banach riflessivo, viceversa dalla (8) consegue che l'operatore A è compatto.

Tutte le nozioni relative agli spazi topologici, e quelle relative agli spazi vettoriali topologici, naturalmente si possono riferire allo spazio S_d, cioè allo spazio vettoriale topologico che si ottiene associando la topologia debole $\tau(S, S^*)$ al sostegno di S (spazio vettoriale normato). Per esprimere che tali nozioni si riferiscono allo spazio S_d, nella relativa terminologia si aggiunge la parola "debolmente", come abbiamo fatto per la convergenza di una successione in S_d. Si parla perciò, con riferimento a sottoinsiemi di S, di insiemi debolmente chiusi, debolmente precompatti, debolmente sequenzialmente compatti, ecc.; si parla anche di chiusura debole, derivato debole ecc.. Tenuto conto della (12.1) di pag. 195, proposiz. 1, è opportuno ricordare quanto si è detto nel n.8 di pag. 24, in particolare a pag. 25.

Se ne deduce ad esempio che un insieme compatto è anche debolmente compatto, e che un insieme debolmente chiuso è anche chiuso.

Segnaliamo due teoremi in proposito, il primo dei quali è noto come teorema di Eberlein-Šmulian.

(12.10). Se S è uno spazio di Banach, per ogni $X \subseteq S$ le proposizioni che seguono sono equivalenti:

a) X è debolmente precompatto.

b) X è debolmente sequenzialmente precompatto.

(12.11). Se S è uno spazio di Banach riflessivo, qualunque sia $X \subseteq S$ le proposizioni che seguono sono equivalenti:

a) X è debolmente compatto.

b) X è limitato e debolmente chiuso.

Notiamo che l'equivalenza della (12.10) estende allo spazio vettoriale topologico S_d l'equivalenza espressa per gli spazi metrici dal teor. (2.1) di pag. 111. Notiamo pure che l'equivalenza della (12.11) è un'estensione della ben nota caratterizzazione dei compatti di R^n.

* * *

Tutte le considerazioni fin quì svolte, essendo riferite ad un qualunque spazio vettoriale normato, sono naturalmente valide se si sostituisce lo spazio S con S^*: si può quindi considerare la topologia debole per S^* che, conformemente alla notazione prima introdotta, si indicherà con

$$\tau(S^*, S^{**}).$$

Associando tale topologia al sostegno di S^* (ossia al duale topologico di S), si ottiene uno spazio vettoriale topologico localmente convesso.

Abbiamo visto che al sostegno di S^* si associa anche la topologia della convergenza puntuale [n.11 pag.191], che abbiamo denotato con $\tau(S^*, S)$. Sarebbe facile verificare che risulta:

$$(9) \qquad \tau(S^*, S) \subseteq \tau(S^*, S^{**}),$$

cioè che la topologia della convergenza puntuale per S^* è più debole della topologia debole per S^*. Rinunciamo alla dimostrazione, segnalando che:

(12.12). <u>Se S è uno spazio di Banach riflessivo, la topologia per S^* della convergenza puntuale e la topologia debole per S^* coincidono.</u>

I simboli da noi adoperati consentono una giustificazione formale della (12.12): se S è uno spazio di Banach riflessivo, lo spazio S^{**} è equivalente ad S, e se si identificano tali spazi i due membri della (9) vengono a coincidere.

Notiamo che per la (12.12) le due topologie coincidono in particolare se S è uno spazio di Hilbert.

<u>Esempio.</u>

Per dare un esempio di utilizzazione dei concetti introdotti in questo numero, proviamo che l'operatore integrale T_1 introdotto a pag.169 è compatto. Ci occorrerà il seguente risultato della teoria dell'integrazione:

(*) <u>Sia $\{y_n\}$ una successione di elementi di $L^2(\Omega)$ convergente ad $y_0 \in L^2(\Omega)$ quasi ovunque in Ω. Se gli integrali di $|y_n|^2$ sono equiassolutamente continui, cioè se</u>

$$\forall \varepsilon > 0 \ \exists \delta > 0 \text{ tale che: } E \subseteq \Omega \text{ e mis. } E < \delta \ (E \text{ misurabile}) \Rightarrow \int_E |y_n(t)|^2 dt < \varepsilon \quad \forall n \in N,$$

<u>allora $\{y_n\}$ converge ad y_0 in $L^2(\Omega)$.</u>

Consideriamo dunque un arbitrario sottoinsieme limitato X di $L^2(\Omega_2)$ e una successione $\{x_n\}$ di punti di X. Per il teor. (12.6) di pag.198 dalla successione $\{x_n\}$ si può estrarre una successione debolmente convergente, che per semplicità denotiamo ancora con $\{x_n\}$, ad un elemento x_0 di $L^2(\Omega_2)$. Si ha allora $\lim (x_n, y) = (x_0, y)$ $\forall y \in L^2(\Omega)$, e quindi in particolare $\lim T_1 x_n = T_1 x_0$ quasi ovunque in Ω.

D'altra parte, posto $y_n = T_1 x_n$ e detto h un numero maggiorante della successione $\{x_n\}$ in $L^2(\Omega)$, per ogni insieme misurabile $E \subseteq \Omega_1$ si ha [cfr. (13) pag.169 con E in luogo di Ω_1]

$$\int_E |y_n(t)|^2 dt \leq h^2 \iint_{E \times \Omega_2} |K(s,t)|^2 ds\, dt$$

e quindi, per l'assoluta continuità dell'integrale di $|K(s,t)|^2$ (cfr. pag.45), gli integrali di $|y_n|^2$ sono equiassolutamente continui. In virtù del teor. (*) sopra enunciato, si ha che $\{y_n\}$ converge in $L^2(\Omega)$ a $T_1 x_0$.

Da quanto sopra consegue che se $\{y_n\}$ è un'arbitraria successione di punti di $T_1(X)$, dalla $\{y_n\}$ si può estrarre una successione convergente: dunque l'insieme $T_1(X)$ è precompatto.

In definitiva T_1 trasforma insiemi limitati in insiemi precompatti: ossia, l'operatore T_1 è compatto.

Capitolo VII

Lo spazio degli operatori lineari e continui. Serie di operatori. Operatore aggiunto. Operatori autoaggiunti.

1. Lo spazio degli operatori lineari e continui di uno spazio vettoriale normato in un altro. Topologia della convergenza forte e topologia della convergenza debole degli operatori.

Siano S, S_1 spazi vettoriali normati, rispettivamente di sostegno Σ e Σ_1. Conformemente alla notazione introdotta nel Cap. III (pag. 69), indichiamo con $\ell(\Sigma, \Sigma_1)$ lo spazio vettoriale costituito dagli operatori lineari di Σ in Σ_1 (*). Il sottoinsieme di $\ell(\Sigma, \Sigma_1)$ costituito dagli operatori lineari di S in S_1 che siano anche continui, è ovviamente un sottospazio vettoriale di $\ell(\Sigma, \Sigma_1)$ che denotiamo con $\ell(S, S_1)$.

Se A è un elemento di $\ell(S, S_1)$, siccome A è lineare e continuo esiste, per il teor. (2.1) di pag. 138, un numero $l \geq 0$ tale da risultare:

$$\|Ax\| \leq l\|x\| \quad \forall x \in S, \tag{1}$$

ed è naturale associare all'operatore A il più piccolo dei numeri l per i quali sussiste la (1), cioè l'estremo superiore dei rapporti:

$$\frac{\|Ax\|}{\|x\|}$$

al variare di x in $S - \{0\}$. Orbene, posto:

$$\nu(A) = \sup_{x \in S - \{0\}} \frac{\|Ax\|}{\|x\|} \tag{2}$$

si definisce un'applicazione ν di $\ell(S, S_1)$ in $[0, +\infty[$ che, come facilmente si verifica, è una norma per $\ell(S, S_1)$.

Lo spazio vettoriale normato $(\ell(S, S_1), \nu)$ si denota col simbolo

(*) Naturalmente Σ e Σ_1 devono essere entrambi reali o entrambi complessi.

$\mathcal{L}(S, S_1)$ e si chiama lo spazio degli operatori lineari e continui di S in S_1.

Si ha dunque, per ogni $A \in \mathcal{L}(S, S_1)$:

$$\|A\| = \sup_{x \in S-\{0\}} \frac{\|Ax\|}{\|x\|} \tag{2'}$$

e quindi:

$$\|Ax\| \leq \|A\| \cdot \|x\| \qquad \forall x \in S. \tag{3}$$

Se S e' reale e $S_1 = R$, gli elementi di $\mathcal{L}(S, S_1)$ sono i funzionali lineari e continui definiti in S: ossia, lo spazio vettoriale $\mathcal{L}(S, R)$ e' il duale topologico di S. In tal caso, per ogni $x \in S$ la norma di Ax in $S_1 = R$ e' il valore assoluto di Ax, e la (2') si riduce alla (4') di pag. 174 (con A in luogo di u): pertanto lo spazio $\mathcal{L}(S, R)$ e' il duale forte S^* dello spazio vettoriale normato S (pag. 173).

Perciò tutte le considerazioni di questo numero, applicate al caso particolare $S_1 = R$, si potranno riferire al duale forte di uno spazio vettoriale normato S reale. Ma ovviamente potranno essere riferite anche al duale forte di uno spazio vettoriale normato S complesso, anche se lo spazio $\mathcal{L}(S, C)$ non coincide con S^* giacché gli elementi di S^* sono i coniugati degli elementi di $\mathcal{L}(S, C)$ [1].

E' importante osservare che qualunque sia $A \in \mathcal{L}(S, S_1)$ risulta:

$$\|A\| = \sup_{x \in S, \|x\| \leq 1} \|Ax\| = \sup_{x \in S, \|x\| = 1} \|Ax\| . \tag{4}$$

Dim. (4). Denotiamo con α e β rispettivamente il secondo e il terzo membro della (4). Per la (3), qualunque sia $x \in S$ tale che $\|x\| \leq 1$ risulta $\|Ax\| \leq \|A\|$: pertanto $\alpha \leq \|A\|$. D'altra parte per ogni $x \in S - \{0\}$, poiché il punto $x/\|x\|$ ha norma 1, si ha:

$$\frac{\|Ax\|}{\|x\|} = \left\| A\left(\frac{x}{\|x\|}\right) \right\| \leq \beta ,$$

e quindi $\|A\| \leq \beta$. Infine e' evidente che $\beta \leq \alpha$. Dunque si ha $\alpha \leq \|A\| \leq \beta \leq \alpha$, donde la (4).

Esempi.

1) Qualunque norma si introduca nello spazio vettoriale $\mathcal{L}(R^k, R^m)$ costituito dagli operatori lineari (e continui) di R^k in R^m [cfr. es. 2 pag. 145], tale norma e' equivalente a quella definita dalla (2). Ciò deriva dal fatto che lo spazio vettoriale $\mathcal{L}(R^k, R^m)$ ha dimensione finita [precisamente

---

(1) Da notare però che, qualunque sia $u \in \mathcal{L}(S, C)$, la norma di u in $\mathcal{L}(S, C)$ e' uguale alla norma di u in S^*.

la sua dimensione è km, come si deduce dall'es.3 di pag. 64] (*).

2) Se S_k, S_m sono due spazi normati a dimensione finita, per lo spazio vettoriale $\mathcal{L}(S_k, S_m)$ vale quanto detto nel precedente esempio con riferimento allo spazio vettoriale $\mathcal{L}(R^k, R^m)$ [si tenga presente l'es.3 di pag. 145].

3) Detto Ω un intervallo limitato di R, l'operatore $A: x \in C^1(\Omega) \to x' \in C^0(\Omega)$ ha norma uguale ad 1.

Essendo banale che $\frac{\|Ax\|_{C^0(\Omega)}}{\|x\|_{C^1(\Omega)}} \leq 1 \; \forall x \in C^1(\Omega) - \{0\}$, basta verificare [per la definizione di $\|A\|$] che per ogni $\varepsilon \in]0,1[$ esiste un $x_\varepsilon \in C^1(\Omega)$ tale che $\frac{\|Ax_\varepsilon\|_{C^0(\Omega)}}{\|x_\varepsilon\|_{C^1(\Omega)}} > 1-\varepsilon$. Esplicitando tale disuguaglianza, si è condotti alla costruzione di una funzione $x_\varepsilon \in C^1(\Omega)$ tale che $\sup|x_\varepsilon(t)| < \frac{\varepsilon}{1-\varepsilon} \sup|x_\varepsilon'(t)|$: si comprende che una funzione siffatta esiste, pensando ad es. ad una funzione avente coefficiente continuato in $[0,1]$ che in un punto ha derivata maggiore di un numero prefissato.

4) È possibile dimostrare che nello spazio $\mathcal{L}(H_0^m(\Omega), L^2(\Omega))$ l'operatore:

$$A : x \in H_0^m(\Omega) \to D^\nu x \in L^2(\Omega) \qquad (0 \leq |\nu| \leq m)$$

ha norma uguale ad 1.

5) Se $H \neq \{0\}$ è un sottospazio chiuso di uno spazio di Hilbert S, il proiettore P_H di S su H è un operatore lineare e continuo [cfr. (4.7) pag. 121 e formula (18) pag. 120]. Si ha:

$$\|P_H\| = 1.$$

Infatti dalla formula citata consegue che $\|P_H\| \leq 1$, e d'altra parte $\forall x \in H$ si ha $\|P_H x\| = \|x\|$.

6) Sia H il sottospazio di $\mathcal{L}(C^1(A), C^0(A))$, dove A è una parallelotopia di R^n, costituito dagli operatori lineari e continui A del tipo considerato nell'es.5 di pag. 146.

Adoperando le notazioni ivi introdotte, se ad ogni $A \in H$ si associa il numero $k = \max\{\alpha, \beta, \gamma\}$ si definisce una norma per H, diciamola ν_H.

Si può dimostrare che la norma ν_H è equivalente a quella indotta su H dalla norma (2) di pag. 203.

7) Detti T_1, T_2 gli operatori integrali generati dal nucleo K introdotti nell'es.1 di pag. 168, ovviamente risulta $\|T_1\| \leq \|K\|_2$, $\|T_2\| \leq \|K\|_{L^2}$; in generale non sussiste l'uguaglianza [si può dimostrare che esistono nuclei K per i quali vale il segno $<$].

(*) Se Σ è uno spazio vettoriale a dimensione finita, due norme qualunque ν e ν' per Σ sono equivalenti; in altri termini, l'identità in Σ è un isomorfismo dello spazio normato (Σ, ν) sullo spazio normato (Σ, ν'). Basta infatti ricordare che ogni operatore lineare tra spazi normati a dimensione finita è continuo [es.3) pag. 145].

Torniamo ora ad occuparci in generale dello spazio $\mathcal{L}(S, S_1)$.

Avvertiamo che, tranne nei casi in cui vi è possibilità di equivoci, col simbolo $\mathcal{L}(S,S_1)$ denotiamo sia lo spazio vettoriale normato $(\mathcal{L}(S,S_1), \nu)$, sia il sostegno di esso.

Osserviamo in primo luogo che:

(1.1) Per ogni successione $(A_n)_{n\in N}$ di elementi di $\mathcal{L}(S,S_1)$, e per ogni $A_0 \in \mathcal{L}(S,S_1)$, le seguenti proposizioni sono equivalenti:

a) La successione $(A_n)_{n\in N}$ converge ad A_0.

b) Qualunque sia X sottoinsieme limitato di S, la successione di operatori $(A_n)_{n\in N}$ converge uniformemente in X ad A_0 (1).

Dim. a)$\Rightarrow$b). Detto X un sottoinsieme limitato di S, sia $K>0$ tale da risultare $\|x\| \le K \quad \forall x \in X$. In virtù di a), per ogni $\varepsilon > 0$ esiste $\nu \in N$ tale che:

$$\|A_n - A_0\| < \varepsilon/K \quad \forall n > \nu .$$

Pertanto, per ogni $n > \nu$ e per ogni $x \in X$ si ha, tenendo presente la (3):

$$\|A_n x - A_0 x\| = \|(A_n - A_0)x\| \le \|A_n - A_0\| \cdot \|x\| \le \frac{\varepsilon}{K} \|x\| < \varepsilon$$

e con ciò la b) è dimostrata.

Dim. b)$\Rightarrow$a). Detto X l'insieme degli $x \in S$ tali che $\|x\| \le 1$, per ogni $\varepsilon > 0$ esiste $\nu \in N$ tale che:

$$\|A_n x - A_0 x\| < \frac{\varepsilon}{2} \quad \forall n > \nu \text{ e } \forall x \in X ;$$

conseguentemente:

$$\|A_n - A_0\| = \sup_{x\in X} \|(A_n - A_0)x\| < \varepsilon \quad \forall n > \nu .$$

Pertanto è vera la a).

(1.2) Nell'ipotesi S_1 spazio di Banach, sia $(A_n)_{n\in N}$ una successione di elementi di $\mathcal{L}(S,S_1)$ convergente ad A_0.

Se per ogni $n \in N$ l'operatore A_n è compatto, allora anche A_0 è compatto.

---

(1) Cioè: per ogni $\varepsilon > 0$ esiste $\nu \in N$ tale che:

$$\|A_n x - A_0 x\| < \varepsilon \quad \forall n > \nu \text{ e } \forall x \in X .$$

Dim. Detto X un sottoinsieme limitato di S, si tratta di far vedere che $A_0(X)$ è precompatto, ossia sequenzialmente precompatto. Consideriamo dunque una successione $(A_0 x_n)_{n\in N}$ di punti di $A_0(X)$, e proviamo che da essa si può estrarre una successione convergente.

Poiché A_1 è compatto, dalla successione $(A_1 x_n)_{n\in N}$ si può estrarre una successione convergente, diciamola $(A_1 x_{n_1(n)})_{n\in N}$; analogamente, dalla successione $(A_2 x_{n_1(n)})_{n\in N}$ si può estrarre una successione convergente, diciamola $(A_2 x_{n_2(n)})_{n\in N}$: ovviamente, anche la successione $(A_1 x_{n_2(n)})_{n\in N}$ è convergente. Procedendo induttivamente, per ogni $K>2$ sia $(A_K x_{n_K(n)})_{n\in N}$ una successione convergente estratta da $(A_K x_{n_{K-1}(n)})_{n\in N}$. Consideriamo quindi la successione "diagonale":

$$(5) \qquad (x_{n_n(n)})_{n\in N}$$

che è estratta da $(x_n)_{n\in N}$. Evidentemente, ciascuno degli operatori $A_1,\dots,A_n,\dots$ trasforma la (5) in una successione convergente; mostriamo che anche A_0 trasforma la (5) in una successione convergente.

Detto ℓ un numero positivo tale che $\|x_n\| \le \ell \ \forall n\in N$, fissato $\varepsilon>0$ scegliamo $K\in N$ in modo che:

$$(6) \qquad \|A - A_K\| < \varepsilon/3\ell \ .$$

Poiché A_K trasforma la (5) in una successione convergente, esiste $\nu\in N$ tale che:

$$(6') \qquad \|A_K x_{n_p(p)} - A_K x_{n_q(q)}\| < \varepsilon/3\ell \quad \forall p,q > \nu \ .$$

Tenendo conto di (6), (6'), per ogni $p>\nu$ e $\forall q>\nu$ si ha:

$$\|A_0 x_{n_p(p)} - A_0 x_{n_q(q)}\| \le$$

$$\le \|A_0 x_{n_p(p)} - A_K x_{n_p(p)}\| + \|A_K x_{n_p(p)} - A_K x_{n_q(q)}\| + \|A_K x_{n_q(q)} - A_0 x_{n_q(q)}\| <$$

$$< \|A_0 - A_K\| \cdot \|x_{n_p(p)}\| + \varepsilon/3\ell + \|A_K - A_0\| \cdot \|x_{n_1(q)}\| <$$

$$< \|A_0 - A_K\| \cdot \ell + \varepsilon/3\ell + \|A_K - A_0\| \cdot \ell < \varepsilon \ .$$

Dunque la successione:

$$(A_0 x_{n_n(n)})_{n\in N}$$

è una successione di Cauchy, e pertanto è convergente giacché S_1 è completo.

La proposizione seguente è una generalizzazione della (6.2) a pag. 174, e si dimostra con lo stesso ragionamento : basta ivi sostituire il modulo dei numeri che vi figurano con la norma in S_1 .

(1.3) Se S_1 è uno spazio di Banach, anche $\mathcal{L}(S,S_1)$ è uno spazio di Banach.

Fondamentale è il seguente teorema (principio di uniforme limitatezza) :

(1.4) Nell'ipotesi S spazio di Banach, sia $(A_n)_{n\in N}$ una successione di elementi di $\mathcal{L}(S,S_1)$.

Se per ogni $x\in S$ la successione di punti di S_1 :

$$(A_n x)_{n\in N}$$

è limitata, allora anche la successione $(A_n)_{n\in N}$ è limitata.

Dim. La dimostrazione è la stessa del teor. (11.2) di pag. 192. Quanto si riferisce ad una successione (u_n) di elementi di S^X , che si suppone puntualmente convergente : ma nel ragionamento la convergenza di $(u_n(x))_{n\in N}$ $\forall x\in S$ viene utilizzata solo per il fatto che tale successione è limitata per ogni $x\in S$.

Naturalmente, considerando la successione di operatori $(A_n)_{n\in N}$ in luogo di $(u_n)_{n\in N}$, bisognerà sostituire il modulo di $u_n(x)$, ove esso figura, con la norma in S_1 si Ax .

Sia $(A_n)_{n\in N}$ una successione di elementi di $\mathcal{L}(S,S_1)$. Si dice che la successione $(A_n)_{n\in N}$ è convergente puntualmente (in S) in senso forte [in senso debole], quando per ogni $x\in S$ la successione di punti di S_1 :

(7) $$(A_n x)_{n\in N}$$

è convergente fortemente [debolmente].

Se $(A_n)_{n\in N}$ è convergente puntualmente in senso forte, si può considerare l'operatore:

$$A_0 : x \in S \to (\lim A_n x) \in S_1 \tag{8}$$

che è ovviamente lineare, e si chiama il limite puntuale forte della successione $(A_n)_{n\in N}$.

Se $(A_n)_{n\in N}$ è convergente puntualmente in senso debole, non è detto che per ogni $x\in S$ la successione (7) sia dotata di limite debole; quando tale circostanza si verifica, si può considerare l'operatore lineare (8) con il limite inteso in senso debole: si dice allora che la successione $(A_n)_{n\in N}$ è dotata di limite puntuale debole, e l'operatore (8) nel senso predetto si chiama il limite puntuale debole di $(A_n)_{n\in N}$.

Sussiste il teorema:

(1.5). Sia $(A_n)_{n\in N}$ una successione di elementi di $\mathcal{L}(S, S_1)$.

Supposto S spazio di Banach, se $(A_n)_{n\in N}$ è convergente puntualmente in senso forte, il limite puntuale forte di $(A_n)_{n\in N}$ è un elemento di $\mathcal{L}(S, S_1)$.

Nell'ulteriore ipotesi che S_1 sia uno spazio di Banach riflessivo, se $(A_n)_{n\in N}$ è convergente puntualmente in senso debole, allora $(A_n)_{n\in N}$ è dotata di limite puntuale debole, e quest'ultimo appartiene ad $\mathcal{L}(S, S_1)$.

Dim. Se $(A_n)_{n\in N}$ è convergente puntualmente in senso forte, per ogni $x\in S$ la successione (7) è limitata; di conseguenza, per la (1.4), la successione $(A_n)_{n\in N}$ è limitata, ossia esiste $l>0$ tale da risultare $\|A_n\|\le l$ $\forall n\in N$, e ciò implica, per la (3):

$$\|A_n x\| \le l\|x\| \quad \forall x\in S \text{ e } \forall n\in N, \tag{9}$$

da cui:

$$\|A_0 x\| \le l\|x\| \quad \forall x\in S. \tag{10}$$

Dunque A_0 è continuo, e pertanto è un elemento di $\mathcal{L}(S, S_1)$.

Nell'ulteriore ipotesi formulata, in virtù della (12.4) di pag. 197 per ogni $x\in S$ la successione $(A_n x)_{n\in N}$ è dotata di limite debole, e quindi $(A_n)_{n\in N}$ è dotata di limite puntuale debole, diciamolo A_0.

Siccome per ogni $x\in S$ la successione (7) è limitata [cfr. (12.5) pag. 198], ancora per la (1.4) sussiste la (9). Ma dalla (9) non si può dedurre la (10), perché non è detto che la successione $(\|A_n x\|)_{n\in N}$ sia convergente.

Fissato $x\in S$ tale che $A_0 x \neq 0$, applichiamo allora la (9.3) di pag. 184. Se ne deduce che esiste $u\in S_1^*$, di norma unitaria, tale da risultare:

$$|\langle u, A_0 x\rangle| = \|A_0 x\|,$$

ossia, per la definizione di convergenza debole:

$$\|A_0 x\| = \lim |\langle u, A_n x\rangle|. \tag{11}$$

D'altra parte si ha, tenuto conto della (9):

$$|\langle u, A_n x\rangle| \le \|u\|\cdot\|A_n x\| = \|A_n x\| \le l\|x\|,$$

e quindi dalla (11) segue la (10). Poiché la (10) sussiste banalmente se $A_0 x = 0$, si conclude che A_0 è continuo e l'asserto è dimostrato.

Il fatto che la successione $(A_n)_{n\in N}$ di elementi di $\mathcal{L}(S, S_1)$ converga puntualmente in senso forte [in senso debole] ad un elemento A_0 di $\mathcal{L}(S, S_1)$, esprime l'ordinaria convergenza di $(A_n)_{n\in N}$ ad A_0 in uno spazio vettoriale topologico che si ottiene associando ad $\mathcal{L}(S, S_1)$ una certa topologia θ_1 [θ_2], che si chiama la <u>topologia della convergenza puntuale forte [debole] per $\mathcal{L}(S, S_1)$</u>, oppure la <u>topologia della convergenza forte [debole] degli operatori</u>.

Per brevità rinunciamo alla definizione delle suddette topologie; ci limitiamo a dichiarare quanto segue: i corrispondenti spazi vettoriali topologici sono spazi di Hausdorff e sono localmente convessi; inoltre la topologia della convergenza puntuale debole θ_2 è più debole della topologia della convergenza puntuale forte θ_1, e questa ultima è più debole della topologia generata dalla norma per $\mathcal{L}(S, S_1)$. Se S_1 è a dimensione finita, le topologie θ_1 e θ_2 coincidono; in particolare nel caso $S_1 = K$, caso in cui gli elementi di $\mathcal{L}(S, S_1)$ sono i coniugati degli elementi di S^*, le topologie θ_1 e θ_2 si possono riferire agli elementi di S^*, e coincidono entrambe con la topologia della convergenza puntuale per S^*. E' questo il motivo per cui, con riferimento ad una successione di elementi di S^*, si parla di convergenza puntuale senza aggiungere l'attributo "forte"o"debole".

Per esprimere che una successione $(A_n)_{n\in N}$ di elementi di $\mathcal{L}(S, S_1)$ è convergente nello spazio $\mathcal{L}(S, S_1)$, cioè rispetto alla topologia generata dalla norma in $\mathcal{L}(S, S_1)$, si dice anche che $(A_n)_{n\in N}$ è <u>convergente in norma</u>.

Naturalmente, se la successione $(A_n)_{n\in N}$ converge in norma ad A_0, essa converge ad A_0 anche puntualmente in senso forte; inoltre, se si verifica quest'ultima circostanza, la successione converge ad A_0 anche puntualmente in senso debole.

Concludiamo questo numero con le seguenti ovvie osservazioni.

1) Qualunque siano gli operatori:

$$A \in \mathcal{L}(S, S_1) \quad , \quad B \in \mathcal{L}(S_1, S_2)$$

con S, S_1, S_2 spazi vettoriali normati, l'operatore $BA \in \mathcal{L}(S, S_2)$ verifica la relazione:

(12) $$\|BA\| \leq \|B\| \cdot \|A\| \ .$$

Dim. Qualunque sia $x \in S$ si ha, tenendo presente la (3):

$$\|(BA)x\| = \|B(Ax)\| \leq \|B\| \cdot \|Ax\| \leq \|B\| \cdot \|A\| \cdot \|x\|$$

da cui la (12).

---

2) Qualunque siano le successioni di operatori $(A_n)_{n\in N}$, $(B_n)_{n\in N}$, con

$$A_n \in \mathcal{L}(S,S_1) \quad , \quad B_n \in \mathcal{L}(S_1,S_2)$$

e S, S_1, S_2 spazi vettoriali normati, dalle relazioni:

$$\lim_{n\to\infty} A_n = A \quad , \quad \lim_{n\to\infty} B_n = B$$

consegue:

(13) $$\lim_{n\to\infty} B_n A_n = BA \quad .$$

Dim. Basta osservare che risulta, per ogni $n\in N$:

$$\|B_n A_n - BA\| \le \|B_n A_n - B_n A\| + \|B_n A - BA\| = \|B_n (A_n - A)\| + \|(B_n - B)A\|$$

e che pertanto, in virtù di (12): $\|B_n A_n - BA\| \le \|B_n\| \cdot \|A_n - A\| + \|B_n - B\| \cdot \|A\|$.

---

2. Applicazioni bilineari continue. Forme sesquilineari continue.

Siano S, S_1, S_2 spazi vettoriali normati . Per ogni applicazione :

(1) $$b : S_1 \times S_2 \to S$$

e per ogni $(x_0,y_0)\in S_1\times S_1$, denotiamo con $b(\cdot,y_0)$ e con $b(x_0,\cdot)$ rispettivamente gli operatori :

(2) $$x\in S_1 \longrightarrow b(x,y_0)\in S \quad , \qquad y\in S_2 \to b(x_0,y)\in S \quad .$$

Diremo che la (1) è un'applicazione bilineare (di $S_1\times S_2$ in S), quando per ogni $(x_0,y_0)\in S_1\times S_2$ gli operatori (2) sono lineari.

Le applicazioni bilineari di $S_1\times S_2$ in R sono le forme bilineari in $S_1\times S_2$ introdotte a pag. 67.

Osserviamo che:

(2.1) Nell'ipotesi che almeno uno degli spazi S_1, S_2 sia uno spazio di Banach, per ogni applicazione bilineare (1) le seguenti proposizioni sono equivalenti :

a) La (1) è continua

b) Esiste $\ell > 0$ tale che:

(3) $$\|b(x,y)\| \le \ell \|x\| \cdot \|y\| \qquad \forall (x,y)\in S_1\times S_2 \quad .$$

Dim. a) $\Rightarrow$ b) . Esaminiamo il caso S_1 spazio di Banach; nel caso S_2 spazio di Banach il ragionamento è del tutto analogo. Posto:

$$\Gamma = \{y\in S_2 : \|y\| \le 1\} \quad ,$$

per ogni $y\in\Gamma$ denotiamo con A_y l'operatore lineare e continuo $b(\cdot,y)$. Si ha allora, qualunque sia $x\in S_1$:

$$(4) \qquad \|A_y x\| = \|b(x,y)\| \le \|b(x,\cdot)\| \cdot \|y\| \le \|b(x,\cdot)\| \ .$$

Dimostriamo che esiste $\ell > 0$ tale che:

$$(5) \qquad \|A_y\| \le \ell \quad \forall y \in \Gamma \ .$$

Se così non fosse, esisterebbe una successione $(y_n)_{n\in N}$ di punti di Γ tale da risultare:

$$(6) \qquad \|A_{y_n}\| > n \quad \forall n \in N \ .$$

Poiché per ogni $x \in S_1$ la successione $(A_{y_n} x)_{n\in N}$ è limitata in virtù della (4), per il principio di uniforme limitatezza [cfr. (1.4)] esiste $h > 0$ tale che:

$$\|A_{y_n}\| \le h \quad \forall n \in N$$

e ciò contraddice la (6).

Tenendo presente la (5) si ha, qualunque sia $x \in S_1$:

$$\|b(x,\cdot)\| = \sup_{\|y\|\le 1} \|b(x,y)\| = \sup_{y\in\Gamma} \|A_y x\| \le \sup_{y\in\Gamma} \|A_y\| \cdot \|x\| \le \ell \|x\| \ .$$

Conseguentemente qualunque siano $x \in S_1$ e $y \in S_2$:

$$\|b(x,y)\| \le \|b(x,\cdot)\| \cdot \|y\| \le \ell \|x\| \cdot \|y\| \ .$$

e la b) è dimostrata.

Dim. b) $\Rightarrow$ a). Fissato $(x_0,y_0) \in S_1 \times S_2$, per ogni $\varepsilon > 0$ sia δ un numero positivo tale che:

$$\delta < 1 \quad , \quad \delta < \frac{\varepsilon}{\ell(\|x_0\| + \|y_0\| + 1)} \ .$$

Qualunque sia $(x,y) \in S_1 \times S_2$ tale che:

$$\|x - x_0\| < \delta \quad , \quad \|y - y_0\| < \delta$$

si ha allora:

$$\|b(x,y) - b(x_0,y_0)\| \le \|b(x,y) - b(x_0,y)\| + \|b(x_0,y) - b(x_0,y_0)\| =$$
$$= \|b(x-x_0,y)\| + \|b(x_0,y-y_0)\| \le \ell \|x-x_0\| \cdot \|y\| + \ell \|x_0\| \cdot \|y-y_0\| <$$
$$< \ell\delta(\|y_0\| + \delta) + \ell \|x_0\| \delta < \ell\delta(\|x_0\| + \|y_0\| + 1) < \varepsilon$$

e dopo ciò la a) è acquisita.

Denotiamo ora con $\mathcal{B}(S_1 \times S_2, S)$ lo spazio vettoriale delle applicazioni bilineari e continue di $S_1 \times S_2$ in S, supponendo che almeno uno degli spazi S_1, S_2 sia completo.

Posto, per ogni $b \in \mathcal{B}(S_1 \times S_2, S)$:

$$\nu(b) = \sup_{\substack{x\in S_1 - \{0\} \\ y \in S_2 - \{0\}}} \frac{\|b(x,y)\|}{\|x\| \cdot \|y\|} \qquad (1)$$

si definisce un'applicazione ν di $\mathcal{B}(S_1 \times S_2, S)$ in $[0, +\infty[$ che, come facilmente si verifica, è una norma per $\mathcal{B}(S_1 \times S_2, S)$.

(1) L'estremo superiore è finito in virtù della (2.1).

Nello spazio vettoriale normato $(\mathcal{B}(S_1\times S_2, S), \nu)$, che denoteremo semplicemente con $\mathcal{B}(S_1\times S_2, S)$ si ha dunque :

$$(7)\qquad \|b\| = \sup_{\substack{x\in S_1-\{0\}\\ y\in S_2-\{0\}}} \frac{\|b(x,y)\|}{\|x\|\cdot\|y\|} \quad ,$$

nonché:

$$(8)\qquad \|b(x,y)\| \leq \|b\|\cdot\|x\|\cdot\|y\| \qquad \forall (x,y)\in S_1\times S_2 \ .$$

Osservazione.

Se gli spazi S_1, S_2 sono complessi, nel caso particolare $S=C$ la (2.1) sussiste anche se l'applicazione (1) è una forma sesquilineare in $S_1\times S_2$. In tal caso si definisce ugualmente la norma di b [cfr.(7)] , e sussiste la (8) .

Consideriamo ora due spazi vettoriali normati X, Y sul campo reale.
Sussiste la proposizione:

(2.2) Se X è uno spazio di Banach, lo spazio vettoriale normato:

$$\mathcal{B}(X\times X, Y)$$

è equivalente allo spazio degli operatori lineari e continui di X in $\mathcal{L}(X,Y)$.

Dim. Per ciascun operatore lineare e continuo A di X in $\mathcal{L}(X,Y)$, poniamo:

$$(9)\qquad b(x,x') = (Ax)x' \qquad \forall (x,x')\in X\times X \ .$$

Evidentemente b è un'applicazione bilineare di $X\times X$ in Y; inoltre, usufruendo:

$$(10)\quad \|b(x,x')\| = \|(Ax)x'\| \leq \|Ax\|\cdot\|x'\| \leq \|A\|\cdot\|x\|\cdot\|x'\| \quad \forall (x,x')\in X\times X$$

in virtù della (2.1) la b è continua . Pertanto $b\in\mathcal{B}(X\times X, Y)$.

Abbiamo così definito un'applicazione, diciamola F , dello spazio

$$(11)\qquad \mathcal{L}(X, \mathcal{L}(X,Y))$$

nello spazio $\mathcal{B}(X\times X, Y)$. Ovviamente F è lineare , e per ogni elemento A di (11) risulta, tenendo presente la (10):

$$(12)\qquad \|F(A)\| \leq \|A\| \ ,$$

D'altro canto, qualunque sia $b\in\mathcal{B}(X\times X, Y)$, ponendo:

$$(13)\qquad Ax = b(x,\cdot) \qquad \forall x\in X$$

si definisce un operatore A di X in $\mathcal{L}(X,Y)$ che è lineare e continuo, e risulta

$b = F(A)$. Pertanto F ha per codominio $\mathcal{B}(X \times X, Y)$, e dalla (13) segue, per ogni $x \in X$:

$$\|Ax\| = \|b(x,\cdot)\| = \sup_{\|x'\| \le 1} \|b(x,x')\| \le \sup_{\|x'\| \le 1} \|b\| \cdot \|x\| \cdot \|x'\| \le \|b\| \cdot \|x\|$$

ossia:

$$\|Ax\| \le \|F(A)\| \cdot \|x\| \qquad \forall x \in X$$

e quindi:

$$\|A\| \le \|F(A)\| .$$

Da tale relazione e dalla (12) consegue $\|F(A)\| = \|A\|$, e quindi F è un isomorfismo isometrico dello spazio (11) su $\mathcal{B}(X \times X, Y)$.

Tenuto conto dell'osservazione di pag. 211, con lo stesso procedimento seguito per il teor. (2.2) si stabilisce il seguente [ponendo $X = S$ spazio di Banach e $Y = K$ (1)]

(2.3) Sia S uno spazio di Banach. Per ogni operatore $A : S \to S^*$, lineare e continuo, la forma sesquilineare in $S \times S$ (associata ad A):

$$b(x,y) = \langle Ax, y \rangle \tag{14}$$

è continua, e risulta $\|b\| = \|A\|$.

Viceversa per ogni forma sesquilineare continua b in $S \times S$, esiste uno ed un solo operatore $A : S \to S^*$, lineare e continuo, tale da verificare la (14).

Se S è uno spazio di Hilbert, utilizzando l'isomorfismo canonico di S su S^* dal teor. (2.3) si deduce il seguente:

(2.4) Sia S uno spazio di Hilbert. Per ogni operatore $A : S \to S$, lineare e continuo, la forma sesquilineare in $S \times S$ (associata ad A):

$$b(x,y) = (Ax, y) \tag{15}$$

è continua, e risulta $\|b\| = \|A\|$.

Viceversa per ogni forma sesquilineare continua b in $S \times S$, esiste uno ed un solo operatore $A : S \to S$, lineare e continuo, tale da verificare la (15).

Del teorema (2.4) vogliamo dare una dimostrazione diretta.

---

(1) Lo spazio $\mathcal{L}(S,K)$ è il duale di S se $K = R$, ma non per $K = C$: peraltro nel ragionamento si può in ogni caso sostituire $\mathcal{L}(S,K)$ con S^*. Notiamo che a secondo membro della (9), essendo Ax un funzionale $\forall x \in S$, subentra $\langle Ax, x' \rangle$.

212.1

Risultando:

(16) $$|b(x,y)| = |(Ax,y)| \le \|Ax\| \cdot \|y\| \le \|A\| \cdot \|x\| \cdot \|y\|,$$

per la (2.1) la $b(x,y)$ è continua. Dalla (16) si deduce poi che $\|b\| \le \|A\|$.

Al fine di stabilire che $\|b\| = \|A\|$, fissato $x \in S$ e detto j l'isomorfismo canonico di S su S^* [pag.179], poniamo:

$$jAx = u.$$

Si ha, tenuto conto della (15):

$$u(y) = \langle jAx, y \rangle = (Ax,y) = b(x,y) \qquad \forall y \in S$$

e quindi:

$$\|Ax\| = \|jAx\| = \|u\| = \sup_y \frac{|u(y)|}{\|y\|} = \sup_y \frac{|b(x,y)|}{\|y\|}$$

$$\le \sup_y \frac{\|b\| \cdot \|x\| \cdot \|y\|}{\|y\|} = \|b\| \cdot \|x\|.$$

Dividendo per $\|x\|$, ne segue che $\|A\| \le \|b\|$, donde l'asserto.

Per provare il viceversa dell'enunciato, sia data una forma bilineare continua $b(x,y)$.

Siccome per ogni $x \in S$ si ha $b(x,\cdot) \in S^*$, per il teorema di Riesz [(7.1) pag. 177] esiste $w_x \in S$ tale che:

(17) $$b(x,\cdot) = (w_x, \cdot).$$

Consideriamo allora l'operatore:

(18) $$A : x \in S \to w_x \in S$$

ovviamente lineare. Risultando:

$$\|Ax\| = \|w_x\| = \|b(x,\cdot)\| = \sup_y \frac{|b(x,y)|}{\|y\|} \le \|b\| \cdot \|x\|,$$

l'operatore A è anche continuo.

Per le (18), (17) si ha poi:

$$\langle Ax, y \rangle = (w_x, y) = b(x,y)$$

e quindi A soddisfa alla condizione (15).

La (2.4) è così dimostrata.

Concludiamo questo numero introducendo per le forme bilineari in uno spazio di Hilbert reale la nozione di coercitivita'.

Sia S uno spazio di Hilbert reale, e sia $b(x,y)$ una forma bilineare in $S \times S$.

Notiamo che se $b(x,y)$ e' continua, per la (8) risulta:

$$b(x,x) \leqslant \|b\| \cdot \|x\|^2 \qquad \forall x \in S.$$

Orbene, la proprieta' di coercitivita' esprime il fatto che $b(x,x)$ verifichi una disuguaglianza opposta. Precisamente, si dice che $b(x,y)$ e' <u>coercitiva</u> (o <u>coerciva</u>) quando esiste un numero positivo α tale da risultare:

$$b(x,x) \geqslant \alpha \|x\|^2 \qquad \forall x \in S.$$

<u>Esempi.</u>

1) Fissato $z \in L^\infty(\Omega)$ con Ω intervallo limitato di R, la forma bilineare $b(x,y)$ definita ponendo:

$$b(x,y) = \int_\Omega [x'(t)y'(t) + z(t)x(t)y(t)]\,dt \qquad \forall (x,y) \in H_0^1(\Omega) \times H_0^1(\Omega)$$

e' continua, ed e' coercitiva se $z \geqslant 0$.

Per la continuita' basta osservare che risulta:

$$|b(x,y)| \leqslant \int_\Omega |x'| \cdot |y'|\,dt + \|z\|_{L^\infty(\Omega)} \cdot \int_\Omega |x| \cdot |y|\,dt \leqslant \|x'\|_{L^2(\Omega)} \cdot \|y'\|_{L^2(\Omega)} + \|z\|_{L^\infty(\Omega)} \cdot \|x\|_{L^2(\Omega)} \cdot \|y\|_{L^2(\Omega)} \leqslant$$

$$\leqslant \|x\|_{H_0^1(\Omega)} \cdot \|y\|_{H_0^1(\Omega)} + \|z\|_{L^\infty(\Omega)} \cdot \|x\|_{H_0^1(\Omega)} \cdot \|y\|_{H_0^1(\Omega)} = \left(1 + \|z\|_{L^\infty(\Omega)}\right) \|x\|_{H_0^1(\Omega)} \cdot \|y\|_{H_0^1(\Omega)}.$$

Se $z \geqslant 0$, tenendo presente la disuguaglianza di Friedrics [cfr. (4) pag. 104] si ha, posto $\alpha = \min\left\{\frac{1}{2}, \frac{1}{2}\frac{1}{(mis\,\Omega)^2}\right\}$:

$$b(x,x) = \int_\Omega [x'^2(t) + z(t)x^2(t)]\,dt \geqslant \int_\Omega x'^2(t)\,dt = \frac{1}{2}\int_\Omega x'^2(t)\,dt + \frac{1}{2}\int_\Omega x'^2(t)\,dt \geqslant$$

$$\geqslant \frac{1}{2}\int_\Omega x'^2(t)\,dt + \frac{1}{2}\frac{1}{(mis\,\Omega)^2}\int_\Omega x^2(t)\,dt \geqslant \alpha \|x'\|^2_{L^2(\Omega)} + \alpha \|x\|^2_{L^2(\Omega)} = \alpha \|x\|^2_{H_0^1(\Omega)}$$

e pertanto $b(x,y)$ e' coercitiva.

2) Fissato $z \in L^\infty(\Omega)$ con Ω intervallo limitato di R, la forma bilineare $b(x,y)$ definita ponendo:

$$b(x,y) = \int_\Omega z(t)x''(t)y''(t)\,dt \qquad \forall (x,y) \in H_0^2(\Omega) \times H_0^2(\Omega)$$

e' continua, ed e' coercitiva se $z(t) \geqslant z_0 > 0$ q.o. in Ω (z_0 = costante).

Per la continuità basta osservare che risulta:

$$|b(x,y)| \leq \|z\|_{L^\infty(\Omega)} \cdot \int_\Omega |x''|\cdot|y''|\,dt \leq \|z\|_{L^\infty(\Omega)} \cdot \|x''\|_{L^2(\Omega)} \cdot \|y''\|_{L^2(\Omega)} \leq \|z\|_{L^\infty(\Omega)} \cdot \|x\|_{H_0^2(\Omega)} \cdot \|y\|_{H_0^2(\Omega)}.$$

Se $z(t) \geq z_0 > 0$, per la disuguaglianza di Friedrichs si ha, denotando con ω la misura di Ω:

$$\|x\|^2_{L^2(\Omega)} \leq \omega^2 \|x'\|^2_{L^2(\Omega)}, \quad \|x'\|^2_{L^2(\Omega)} \leq \omega^2 \|x''\|^2_{L^2(\Omega)} \tag{19}$$

e quindi:

$$\|x\|^2_{H_0^2(\Omega)} = \|x\|^2_{L^2(\Omega)} + \|x'\|^2_{L^2(\Omega)} + \|x''\|^2_{L^2(\Omega)} \leq$$

$$\leq \omega^4 \|x''\|^2_{L^2(\Omega)} + \omega^2 \|x''\|^2_{L^2(\Omega)} + \|x''\|^2_{L^2(\Omega)} = (\omega^4 + \omega^2 + 1)\|x''\|^2_{L^2(\Omega)},$$

da cui segue:

$$\|x''\|^2_{L^2(\Omega)} \geq \frac{1}{\omega^4+\omega^2+1} \|x\|^2_{H_0^2(\Omega)}.$$

Tenuto conto di tale disuguaglianza, si ha:

$$b(x,x) = \int_\Omega z(t) x''^2(t)\,dt \geq z_0 \int_\Omega x''^2(t)\,dt = z_0 \|x''\|^2_{L^2(\Omega)} \geq \frac{z_0}{\omega^4+\omega^2+1} \|x\|^2_{H_0^2(\Omega)}$$

e pertanto $b(x,y)$ è coercitiva.

3) Detto Ω un intervallo limitato di $\mathbb{R}$, sia H il sottospazio di $H^2(\Omega)$:

$$H = H^2(\Omega) \cap H_0^1(\Omega)$$

che è uno spazio di Hilbert col prodotto scalare:

$$(x,y)_H = \int_\Omega [xy + x'y' + x''y'']\,dt.$$

Fissato $z \in L^\infty(\Omega)$, la forma bilineare $b(x,y)$ definita ponendo:

$$b(x,y) = \int_\Omega [x'(t)y'(t) + z(t)x''(t)y''(t)]\,dt \qquad \forall (x,y) \in H \times H$$

è continua, ed è coercitiva se $z(t) \geq z_0 > 0$ q.o. in Ω (z_0 = costante).

Per la continuità si procede come negli esempi precedenti. Quanto alla coercitività, tenendo presente la prima delle disuguaglianze (19) dianzi utilizzate si ha:

$$b(x,x) = \int_\Omega x'^2\,dt + \int_\Omega z\,x''^2\,dt \geq \|x'\|^2_{L^2(\Omega)} + z_0 \|x''\|^2_{L^2(\Omega)} = \frac{1}{2}\|x'\|^2_{L^2(\Omega)} + \frac{1}{2}\|x'\|^2_{L^2(\Omega)} + z_0\|x''\|^2_{L^2(\Omega)}$$

$$\geq \frac{1}{2}\|x'\|^2_{L^2(\Omega)} + \frac{1}{2\omega^2}\|x\|^2_{L^2(\Omega)} + z_0 \|x''\|^2_{L^2(\Omega)}$$

da cui, posto $\alpha = \min\left\{\frac{1}{2}, \frac{1}{2\omega^2}, z_0\right\}$, segue:

$$b(x,x) \geq \alpha \|x\|^2_H.$$

3. Serie di operatori.

Siano : S, S_1 spazi vettoriali normati, $(A_n)_{n\in N}$ una successione di operatori lineari e continui di S in S_1.

Consideriamo la serie :

(1) $$A_1 + A_2 + \cdots + A_n + \cdots$$

Posto:

$$B_n = A_1 + A_2 + \cdots + A_n \qquad \forall n \in N .$$

se la successione $(B_n)_{n\in N}$ è convergente (in norma), si dirà che la (1) è convergente (in norma), e il limite di $(B_n)_{n\in N}$, che è un elemento di $\mathcal{L}(S, S_1)$, si chiamerà somma della (1). Analogamente, se la successione $(B_n)_{n\in N}$ è convergente puntualmente in senso forte [risp. debole], si dirà che la serie (1) è convergente puntualmente in senso forte [risp. debole] : con riferimento a tali nozioni, è importante tener presente la proposizione (1.5) a pag. 207.

Osserviamo che:

(3.1). Nell'ipotesi che S_1 sia uno spazio di Banach, se la serie delle norme della (1) è maggiorata da una serie $\sum a_n$ convergente, la (1) è convergente e si ha:

(2) $$\left\| \sum_{n=1}^{\infty} A_n \right\| \leq \sum_{n=1}^{\infty} a_n .$$

Dim. Detta B_n la somma parziale n^{ma} della (1), qualunque siano $n, k \in N$ si ha:

$$\| B_{n+k} - B_n \| = \| A_{n+1} + \cdots + A_{n+k} \| \leq a_{n+1} + \cdots + a_{n+k} .$$

Tenendo presente che $\mathcal{L}(S, S_1)$ è uno spazio di Banach [cfr. (1.3) pag. 206], per la (6.1) di pag. 125 la serie (1) è convergente. Quanto alla (2), si ha:

$$\left\| \sum_{n=1}^{\infty} A_n \right\| = \left\| \lim_{n\to\infty} \sum_{i=1}^{n} A_i \right\| = \lim_{n\to\infty} \left\| \sum_{i=1}^{n} A_i \right\| \leq \lim_{n\to\infty} \sum_{i=1}^{n} \|A_i\| \leq \lim_{n\to\infty} \sum_{i=1}^{n} a_i = \sum_{n=1}^{\infty} a_n .$$

(3.2) Nell'ipotesi S spazio di Banach, siano A un operatore lineare e continuo di S in S, ed I l'identità in S. Se risulta

$$\|A\| < 1 ,$$

la serie

(3) $$I + A + A^2 + \cdots + A^{n-1} + \cdots \qquad (1)$$

è convergente, e si ha

(4) $$\sum_{n=1}^{\infty} A^{n-1} = (I - A)^{-1} \qquad (2).$$

---

(1) Naturalmente per ogni $n \geq 1$ il simbolo A^n denota il prodotto di n operatori uguali ad A (pag. 72). Inoltre si è posto $A^0 = I$.

(2) Precisamente l'operatore $I - A$ è invertibile, e il suo inverso appartiene ad $\mathcal{L}(S, S)$ e verifica la (4).

Dim. Per la (12) di pag. 208 si ha:

$$\|A^n\| \leq \|A\|^n \tag{5}$$

e ovviamente $\|I\| = 1$; d'altra parte essendo $\|A\| < 1$ la serie geometrica:

$$1 + \|A\| + \cdots + \|A\|^{n-1} + \cdots$$

è convergente. Dalla (3.1) consegue allora che la serie (3) è convergente.

Detta B la somma della (3), dalla relazione:

$$(I + A + A^2 + \cdots + A^{n-1})(I - A) = I - A^n$$

consegue, tenendo per la (5) $\lim A_n = 0$:

$$B(I - A) = I .$$

Dunque B è un inverso sinistro di $I - A$, e pertanto $I - A$ è invertibile [cfr. (6) pag. 75].

Analogamente si prova che B è anche un inverso destro di $I - A$, e pertanto $I - A$ ha per codominio tutto lo spazio [cfr. (7) pag. 75]. Ne segue che l'inverso di $I - A$ coincide con B [si tenga presente quanto detto a pag. 74 a proposito della (5)].

(3.3). *Nell'ipotesi S spazio di Banach, siano A un operatore lineare e continuo di S in S, ed I l'identità in S. Qualunque sia $\lambda \in K$ tale che:*

$$|\lambda| > \|A\| ,$$

l'operatore $A - \lambda I$ è invertibile ed ha per codominio tutto lo spazio; l'operatore inverso è continuo (oltre che lineare), ed è la somma della serie:

$$-\frac{I}{\lambda} - \frac{A}{\lambda^2} - \frac{A^2}{\lambda^3} - \cdots - \frac{A^{n-1}}{\lambda^n} - \cdots \tag{6}$$

Dim. Posto $B = A/\lambda$, essendo $\|B\| < 1$ per la (3.2) l'operatore $I - B$ è invertibile ed appartiene ad $\mathcal{L}(S, S)$, risultando inoltre:

$$(I - B)^{-1} = \sum_{n=1}^{\infty} B^{n-1} . \tag{7}$$

Ne segue (1):

$$(\lambda B - \lambda I)^{-1} = \frac{1}{\lambda}(I - B)^{-1} = -\sum_{n=1}^{\infty} \frac{B^{n-1}}{\lambda} = -\sum_{n=1}^{\infty} \frac{A^{n-1}}{\lambda^n} \tag{8}$$

e quindi $(A - \lambda I)^{-1}$, ossia il primo membro della (8), è la somma della serie (6).

---

(1) Ovviamente, se T è un operatore invertibile, qualunque sia $\alpha \in K$ anche l'operatore αT è invertibile [nel nostro caso $T = I - B$ e $\alpha = -\lambda$], e si ha $(\alpha T)^{-1} = \frac{1}{\alpha} T^{-1}$.

Osservazioni.

Vogliamo ritornare sulle definizioni introdotte all'inizio di questo numero [pag. 215] per soffermarci su di esse.

Siano S, S_1 spazi vettoriali normati, $(A_n)_{n\in\mathbb{N}}$ una successione di elementi dello spazio $\mathcal{L}(S,S_1)$ [pag. 203], cioè di operatori lineari e continui di S in S_1, ed A un elemento di $\mathcal{L}(S,S_1)$.

L'uguaglianza

$$\sum_{n=1}^{\infty} A_n = A \tag{9}$$

significa che risulta:

$$\lim_{n\to\infty} \sum_{i=1}^{n} A_i = A \quad \text{in } \mathcal{L}(S,S_1)$$

e ciò equivale all'uguaglianza:

$$\lim_{n\to\infty} \left\| \sum_{i=1}^{n} A_i - A \right\|_{\mathcal{L}(S,S_1)} = 0 .$$

Ricordando quanto detto a pag. 208, dopo la definizione di convergenza in norma, la (9) implica che risulta:

$$\sum_{n=1}^{\infty} A_n x = Ax \quad \forall x \in S \tag{10}$$

e questa significa che:

$$\lim_{n\to\infty} \left\| \sum_{i=1}^{n} A_i x - Ax \right\|_{S_1} = 0 \quad \forall x \in S .$$

E' bene evidenziare che mentre dalla (9) consegue la (10), viceversa la (10) non implica la (9); cioè dall'uguaglianza:

$$A_1 x + A_2 x + \cdots + A_n x + \cdots = Ax \quad \forall x \in S \tag{11}$$

non è lecito dedurre l'altra:

$$A_1 + A_2 + \cdots + A_n + \cdots = A .$$

Aggiungiamo che se è vera la (11), cioè se la serie di operatori $\sum_{n=1}^{\infty} A_n$ converge ad A puntualmente (in senso forte), essa converge puntualmente ad A anche in senso debole; ciò significa che risulta, qualunque sia $x \in S$:

$$\sum_{n=1}^{\infty} \langle u, A_n x \rangle = \langle u, Ax \rangle \quad \forall u \in S_1^* . \tag{12}$$

Viceversa la (12) non implica la (11); dalla (12) si può dedurre la (11) se S, S_1 sono spazi di Banach essendo inoltre S_1 riflessivo [(1.5) pag. 207].

Consigliamo al lettore di ripetere le considerazioni precedenti con riferimento all'uguaglianza (4) di pag. 215 e alla serie (6) di pag. 216.

Esempio

Vogliamo segnalare una importante applicazione del teorema (3.3) di pag. 216.

Detto Ω un aperto limitato di R^k e $K(s,t)$ un elemento di $L^2(\Omega\times\Omega)$, consideriamo l'operatore integrale $J_1 : L^2(\Omega)\to L^2(\Omega)$ generato dal nucleo K con variabile di integrazione t [pag. 169]. Denotato con I l'operatore identico di $L^2(\Omega)$ in $L^2(\Omega)$ e fissati $y\in L^2(\Omega)$, $\lambda\in C$, consideriamo quindi l'equazione integrale nell'incognita $x\in L^2(\Omega)$:

$$(13)\qquad (J_1-\lambda I)x = y$$

che esplicitamente si scrive:

$$\int_\Omega K(s,t)x(t)\,dt - \lambda x(s) = y(s).$$

Orbene, se il numero complesso λ verifica la disuguaglianza:

$$(14)\qquad |\lambda| > \|K\|_{L^2(\Omega\times\Omega)},$$

risultando $\|J_1\|\le\|K\|$ (ciò consegue dalla disuguaglianza che precede la (14) di pag. 169) il numero λ verifica anche l'altra:

$$|\lambda| > \|J_1\|.$$

Pertanto all'operatore J_1 è applicabile il teor. (3.3) di pag. 216. Dunque l'operatore $J_1-\lambda I$ è invertibile ed ha per codominio tutto lo spazio $L^2(\Omega)$, e l'operatore inverso è dato da:

$$(J_1-\lambda I)^{-1} = -\frac{I}{\lambda} - \sum_{i=1}^{\infty}\frac{J_1^i}{\lambda^{i+1}}$$

Siccome la serie a secondo membro converge anche puntualmente in $L^2(\Omega)$ [pag. 217], ne segue che risulta:

$$(15)\qquad (J_1-\lambda I)^{-1}y = \frac{-y}{\lambda} - \sum_{i=1}^{\infty}\frac{J_1^i y}{\lambda^{i+1}}.$$

Si conclude allora che se è verificata la (14), l'equazione (13) ammette qualunque sia $y\in L^2(\Omega)$ una e una sola soluzione, che è l'elemento x_λ di $L^2(\Omega)$ espresso dal secondo membro della (15).

La (13) sarà ripresa nell'Appendice [pag. 335]: vedremo che la soluzione della (13) si può rappresentare mediante un operatore integrale, rappresentato da un opportuno nucleo.

4. Aggiunto hermitiano.

Fissato un numero complesso α, consideriamo l'operatore lineare di C in C:

$$A : x \in C \to \alpha x .$$

Qualunque siano $x, y \in C$ si ha:

$$(Ax, y) = (\alpha x, y) = \alpha (x, y) = (x, \bar{\alpha} y)$$

e quindi l'operatore A^* di C in C:

$$A^* : x \in C \to \bar{\alpha} x$$

verifica l'uguaglianza:

$$(1) \qquad (Ax, y) = (x, A^* y) .$$

Anche per l'operatore lineare:

$$A : x \in H_0^1(R) \to x' \in L^2(R)$$

definito in un sottospazio di $L^2(R)$, esiste un operatore lineare definito nello stesso sottospazio, tale da verificare la (1) qualunque siano $x, y \in H_0^1(R)$, il prodotto scalare essendo inteso in $L^2(R)$. Infatti, essendo:

$$(Ax, y) = \int_R x' y \, dt = - \int_R x y' \, dt \qquad (*),$$

basta indicare con A^* l'operatore che ad ogni $x \in H_0^1(R)$ associa $-x' \in L^2(R)$.

Consideriamo dunque un operatore lineare:

$$A : H \to S ,$$

con S spazio di Hilbert ed H sottospazio vettoriale di S denso in S, e supponiamo che esista un operatore:

$$A^* : H \to S$$

tale da verificare la (1) per ogni $x, y \in H$. L'operatore A^* è necessariamente lineare, ed è univocamente determinato.

— — — — — — — —

Il fatto che A^* è lineare segue dalle relazioni:

$$(A^*(x_1 + x_2), y) = (x_1 + x_2, Ay) = (x_1, Ay) + (x_2, Ay) = (A^* x_1, y) + (A^* x_2, y) = (A^* x_1 + A^* x_2, y)$$

$$(A^*(\lambda x), y) = (\lambda x, Ay) = \lambda (x, Ay) = \lambda (A^* x, y) = (\lambda A^* x, y) .$$

Infatti, essendo:

(*) L'uguaglianza si stabilisce considerando una successione $\{\varphi_n\}$ di elementi di $C_0^\infty(R)$ fondamentale ad y in $H_0^1(R)$ [si ricordi che $C_0^\infty(\Omega)$ è denso in $H_0^{m,p}(\Omega)$, cfr. pag. 104]. Basta allora osservare che per definizione di derivata debole si ha

$$\int_R x' \varphi_n \, dt = - \int_R x \varphi_n' \, dt$$

e passare al limite per $n \to \infty$.

$$(A^*(x_1+x_2)-A^*x_1-A^*x_2\,,\,y)=0 \quad \forall y\in H,$$

la stima uguaglianza sussiste per ogni $y\in S$ in quanto H è denso in S, e ciò implica:

$$A^*(x_1+x_2)-A^*x_1-A^*x_2=0\ ;$$

analogamente si conclude che $A^*(\lambda x)=\lambda A^*x$.

Per provare l'unicità, sia $B:H\to S$ un operatore soddisfacente alla (1) $\forall x,y\in H$, onia tale da risultare:

$$(Ax,y)=(x,By) \quad \forall x,y\in H.$$

Si ha allora:

$$0=(x,(A^*-B)y) \quad \forall x,y\in H$$

e per ogni fissato $y\in H$ l'uguaglianza sussiste qualunque sia $x\in S$, giacchè H è denso in S. Dunque $(A^*-B)y=0$ $\forall y\in H$, ossia $A^*-B=0$: pertanto $B=A^*$

— — — — — — — — — —

Orbene, se esiste un operatore $A^*:H\to S$ tale da risultare:

(2) $$(Ax,y)=(x,A^*y) \quad \forall x,y\in H$$

o equivalentemente:

(3) $$(A^*x,y)=(x,Ay) \quad \forall x,y\in H$$ [1],

A^* si chiama l'operatore <u>aggiunto (hermitiano) di A</u>.

Ovviamente se risulta:

$$(Ax,y)=(x,Ay) \quad \forall x,y\in H,$$

l'aggiunto di A esiste e coincide con A : si dice allora che l'operatore lineare A è <u>autoaggiunto</u>.

Una condizione sufficiente per l'esistenza dell'aggiunto è fornita dal teorema:

(4.1). <u>Sia S uno spazio di Hilbert, e sia H un sottospazio di S il cui sostegno è denso in S.</u>

<u>Per ogni operatore lineare e continuo $A:H\to S$, esiste l'operatore aggiunto A^* ed è anch'esso (lineare e) continuo.</u>

— — — — — — — — — —

Dim. Fissato $x\in H$, consideriamo il funzionale u definito in H ponendo:

(4) $$u(y)=(Ay,x) \quad \forall y\in H,$$

ovviamente lineare. Essendo:

---

(1) Dalla (2) passando ai coniugati si deduce $(y,Ax)=(A^*y,x)$, e questa uguaglianza, scambiando il ruolo delle variabili x ed y, è appunto la (3).

$$|u(y)| \leq \|Ay\| \cdot \|x\| \leq \|A\| \cdot \|x\| \cdot \|y\| \tag{5}$$

il funzionale u è anche continuo. Per il teorema di Hahn-Banach, esiste un prolungamento lineare e continuo $\tilde{u}$ di u su S, univocamente determinato in quanto H è denso in S [1], e si ha: $\|\tilde{u}\|_{S'} = \|u\|_{H'}$. In virtù del teorema di Riesz [pag. 177], esiste poi un unico elemento di S, che dipende dal prefissato $x \in H$ e che perciò denotiamo con w_x, tale che $\tilde{u} = (\cdot, w_x)$, ossia tale che:

$$\tilde{u}(y) = (y, w_x) \quad \forall y \in S,$$

e risulta $\|w_x\|_S = \|\tilde{u}\|_{S'}$. In particolare, essendo $\tilde{u}(y) = u(y)$ $\forall y \in H$, per la (4) si ha:

$$(Ay, x) = (y, w_x) \quad \forall y \in H. \tag{6}$$

Se allora si considera l'operatore:

$$A^* : x \in H \to w_x \in S,$$

si ha $w_x = A^* x$ e la (6) si scrive:

$$(Ay, x) = (y, A^* x).$$

Questa uguaglianza equivale alla (3).

Per provare che A^* è continuo, osserviamo che per la (5) si ha:

$$\|x\|_{H'} \leq \|A\| \cdot \|x\| . \quad \forall x \in H. \tag{7}$$

Essendo $\|u\|_{H'} = \|\tilde{u}\|_{S'} = \|w_x\|_S = \|A^* x\|_S$, la (7) si scrive:

$$\|A^* x\| \leq \|A\| \cdot \|x\| \qquad \forall x \in H \tag{8}$$

e l'asserto è dimostrato.

Nelle osservazioni che seguono denotiamo con A, salvo diversa precisazione, un operatore lineare di H in S, essendo S uno spazio di Hilbert ed H un sottospazio vettoriale di S denso in S [2].

(4.2). Se esiste l'aggiunto di A, esiste anche l'aggiunto dell'aggiunto, e coincide con A.

Si tratta di far vedere che esiste un operatore $(A^*)^*$ tale da risultare:

$$((A^*)^* x, y) = (x, A^* y) . \quad \forall x, y \in H$$

---

(1) Siano: Σ uno spazio topologico di Hausdorff, S_0 un sottoinsieme di Σ denso in Σ. Se f, g sono funzionali continui in Σ, si ha ovviamente:

$$f(x) = g(x) \ \forall x \in S_0 \Rightarrow f = g$$

Basta osservare che per ogni $y \in \Sigma - S_0$ si ha:

$$f(y) = \lim_{\substack{x \to y \\ x \in S_0}} f(x) = \lim_{\substack{x \to y \\ x \in S_0}} g(x) \quad g(y).$$

(2) Quando si supporrà A continuo, naturalmente H si intenderà normato con la norma di S.

e che tale operatore coincide con A. Si tratta cioè di dimostrare che si ha:

$$(Ax, y) = (x, A^* y) \quad \forall x, y \in H ,$$

e ciò è vero per definizione di operatore aggiunto.

(4.3). Se l'operatore lineare A è continuo, la norma dell'aggiunto A^* è uguale alla norma di A.

Invero per la (8) risulta $\|A^*\| \leq \|A\|$, e quindi si ha anche $\|(A^*)^*\| \leq \|A^*\|$, ossia per la (4.2) $\|A\| \leq \|A^*\|$. Ne segue l'asserto.

(4.4). Siano A e B operatori lineari di H in S. Se esistono gli operatori aggiunti A^* e B^*, esiste anche l'aggiunto di $A+B$, e risulta:

$$(9) \qquad (A+B)^* = A^* + B^* \quad ;$$

nell'ipotesi $H = S$ esiste anche l'aggiunto di $A \cdot B$, e risulta:

$$(10) \qquad (AB)^* = B^* A^* .$$

Se esiste l'aggiunto di un operatore lineare A di H in S, per ogni $\lambda \in C$ esiste anche l'aggiunto di λA, e risulta:

$$(\lambda A)^* = \bar{\lambda} A^* .$$

Le precedenti asserzioni si deducono in modo ovvio dalla definizione di aggiunto; dimostriamo ad es. la prima.

Qualunque siano $x, y \in H$ si ha:

$$\big((A+B)x, y\big) = (Ax, y) + (Bx, y) = (x, A^* y) + (x, B^* y) =$$

$$= (x, (A^* + B^*) y) .$$

Ciò significa che esiste l'aggiunto di $A+B$, e che vale la (9).

(4.5). Se A è un isomorfismo di S su S, anche l'aggiunto A^* è un isomorfismo di S su S, e l'inverso dell'aggiunto di A è uguale all'aggiunto dell'inverso di A.

Ovviamente l'operatore identico I di S su S è autoaggiunto. Allora dalle relazioni:

$$A A^{-1} = I \quad , \quad A^{-1} A = I \ ,$$

passando agli aggiunti e tenendo conto della (10) consegue:

$$(A^{-1})^* A^* = I \quad , \quad A^* (A^{-1})^* = I \ .$$

Dunque A^* è dotato di inverso sinistro e di inverso destro. Ne segue che A^* è invertibile ed ha per codominio tutto lo spazio. Di conseguenza l'inverso di A^* è uguale all'inverso destro e sinistro, ossia ad $(A^{-1})^*$.

(4.6). Qualunque sia l'operatore lineare e continuo $A : S \to S$, gli operatori:

$$A A^* \quad , \quad A^* A$$

sono autoaggiunti.

Infatti si ha:

$$(A A^*)^* = A^{**} A^* = A A^* \quad , \qquad (A^* A)^* = A^* A^{**} = A^* A \ .$$

(4.7). Se lo spazio S è complesso, sussiste l'equivalenza:

$$A \text{ autoaggiunto} \iff (Ax, x) \text{ reale } \forall x \in H$$

Dim. Se A è autoaggiunto, per la (4) con $y = x$ si ha:

$$\overline{(Ax, x)} = (x, Ax) = (Ax, x) \quad \forall x \in H$$

e quindi (Ax, x) è reale. Viceversa, se (Ax, x) è reale $\forall x \in H$, qualunque siano $x, y \in H$ risulta:

$$(A(x+y), x+y) = (x+y, A(x+y)) \quad , \quad (A(x+jy), x+jy) = (x+jy, A(x+jy))$$

e queste uguaglianze, tenuto conto ancora dell'ipotesi, equivalgono alle seguenti.

$$(Ay, x) + (Ax, y) = (y, Ax) + (x, Ay) \quad , \quad (Ay, x) - (Ax, y) = (y, Ax) - (x, Ay)$$

da cui, sottraendo membro a membro, segue la (4).

(4.8). Se lo spazio S è complesso, qualunque sia $A : H \to S$ lineare e continuo, esistono due operatori autoaggiunti A_1, A_2 tali da risultare:

$$A = A_1 + j A_2 \quad , \quad A^* = A_1 - j A_2 \ .$$

Basta porre:

$$A_1 = \frac{1}{2}(A + A^*) \quad , \qquad A_2 = \frac{1}{2j}(A - A^*) \ .$$

(4.9). Un operatore autoaggiunto A che verifichi l'uguaglianza:

$$(Ax,x)=0 \quad \forall x\in H \tag{11}$$

è necessariamente l'operatore nullo.

Dim. Proviamo anzitutto che la (11) implica:

$$(Ax,y)=0 \quad \forall x,y\in H. \tag{12}$$

Fissati $x,y\in H$, e posto: $\lambda=(Ax,y)$,

per la (11) risulta:

$$\big(A(x+\lambda y),x+\lambda y\big)=0$$

e questa uguaglianza, ancora per la (11), equivale alla seguente:

$$\bar\lambda(Ax,y)+\lambda(Ay,x)=0. \tag{13}$$

Essendo A autoaggiunto, si ha d'altra parte:

$$(Ay,x)=(y,Ax)=\overline{(Ax,y)}=\bar\lambda$$

e quindi la (13) si scrive:

$$\bar\lambda\lambda+\lambda\bar\lambda=0$$

da cui $|\lambda|^2=0$.

Essendo H denso in S, la (12) sussiste poi qualunque sia $y\in S$, e quindi anche per $y=Ax$. Ne segue $Ax=0$ $\forall x\in H$, ossia $A=0$.

Consideriamo ora un operatore lineare e continuo $A: S\to S$, e supponiamo che sia autoaggiunto.

Anche se S è complesso, per ogni $x\in S$ il prodotto scalare (Ax,x) è reale [cfr. (4.7)], e siccome risulta $|(Ax,x)|\le\|A\|\cdot\|x\|^2$, nell'insieme $\{x\in S:\|x\|=1\}$ si ha $|(Ax,x)|\le\|A\|$.

Pertanto al variare di x nell'insieme predetto il prodotto scalare (Ax,x) descrive un insieme numerico limitato. Orbene, i numeri reali:

$$m(A)=\inf_{\|x\|=1}(Ax,x)=\inf_{x\in S-\{0\}}\frac{(Ax,x)}{\|x\|^2} \tag{14}$$

$$M(A)=\sup_{\|x\|=1}(Ax,x)=\sup_{x\in S-\{0\}}\frac{(Ax,x)}{\|x\|^2} \tag{15}$$

si chiamano gli estremi dell'operatore A. Si dimostra che:

$$\max\{|m(A)|,|M(A)|\}=\|A\|. \tag{16}$$

Se risulta $m(A)\ge 0$, ossia $(Ax,x)\ge 0$ $\forall x\in S$, la forma quadratica (Ax,x) è semidefinita positiva, e perciò si dice che A è un operatore positivo.

Se A e B, operatori lineari e continui di S in S, sono autoaggiunti, si dice che A è maggiore di B, e si scrive $A\ge B$, quando l'operatore $A-B$ è positivo.

Quando risulta $A \geq B$ ed $A \neq B$, si dice che A è strettamente maggiore di B, e si scrive $A > B$.

Osserviamo che:

(4.10). Se l'operatore $A: S \to S$ lineare e continuo è autoaggiunto, allora l'operatore A^2 è positivo.

Invero per ogni $x \in S$ si ha:

$$(A^2x, x) = (A(Ax), x) = (Ax, Ax) = \|Ax\|^2 \geq 0.$$

Il proiettore P_H di S su un sottospazio chiuso di S coincide ovviamente con P_H^2, ed è lineare e continuo [cfr. (4.7) pag. 121 e formula (18) pag. 120]: dunque per la (4.10) P_H è un operatore positivo. Il teorema che segue fornisce una caratterizzazione dei proiettori mediante la nozione di operatore autoaggiunto.

(4.11). Se H è un sottospazio chiuso di uno spazio di Hilbert S, il proiettore P_H di S su H è un operatore autoaggiunto tale che $P_H^2 = P_H$.

Viceversa ogni operatore lineare $A: S \to S$, autoaggiunto e tale che $A^2 = A$, è un proiettore: precisamente, il sottospazio $H = A(S)$ è chiuso e risulta $A = P_H$.

DIM. Detto L il supplemento ortogonale di H [pag. 120], qualunque siano $x, y \in S$ si ha, tenendo presente la (16) di pag. 119:

$$(P_H x, y) = (P_H x, P_H y + P_L y) = (P_H x, P_H y) + (P_H x, P_L y) = (P_H x, P_H y)$$

$$(x, P_H y) = (P_H x + P_L x, P_H y) = (P_H x, P_H y) + (P_L x, P_H y) = (P_H x, P_H y)$$

e quindi P_H è autoaggiunto. Abbiamo già osservato che $P_H^2 = P_H$.

Viceversa, sia $A: S \to S$ lineare, autoaggiunto e tale che $A^2 = A$. Dalla disuguaglianza:

$$\|Ax\|^2 = (Ax, Ax) = (A^2x, x) = (Ax, x) \leq \|Ax\| \cdot \|x\| \quad \forall x \in S$$

consegue che $\|Ax\| \leq \|x\|$ $\forall x \in S$, e quindi che A è continuo. Dimostriamo

che $A(S)$ è chiuso.

Si tratta di stabilire, in virtù della (1.8) di pag.109, che se $(y_n)_{n\in N}$ è una qualunque successione convergente di punti di $A(S)$, il suo limite y_0 appartiene ad $A(S)$. Orbene, posto $y_n = Ax_n \ \forall n\in N$, tenendo presente che A è continuo e che $A^2 = A$, risulta:

$$y_0 = \lim y_n = \lim Ax_n = \lim A^2 x_n = \lim A(Ax_n) = A(\lim Ax_n) = A(\lim y_n) = Ay_0 .$$

Dall'uguaglianza $y_0 = Ay_0$ consegue che $y_0 \in A(S)$.

Detto H il sottospazio chiuso $A(S)$ dello spazio S, dimostriamo infine che per ogni $x\in S$ il punto Ax è la proiezione di x su H. Per ogni $x\in H$, posto $x = Az$ si ha:

$$(x-Ax, h) = (x-Ax, Az) = (x, Az) - (Ax, Az) =$$
$$= (x, Az) - (x, A^2 z) = (x, Az) - (x, Az) = 0 .$$

Risultando $(x-Ax, h) = 0 \ \forall h\in H$, per la (4.3) di pag.118 si ha $Ax = P_H x$.

Il teorema è così dimostrato.

Concludiamo questo numero con un importante teorema, che fornisce un legame tra il codominio di un operatore di S in S e il nucleo dell'operatore aggiunto; enunceremo poi una notevole conseguenza del teorema.

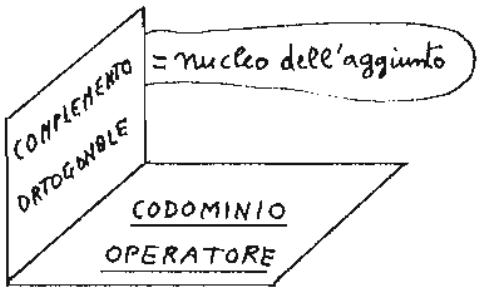
Come al solito S denoterà uno spazio di Hilbert. Ricordiamo che il complemento ortogonale $Y^\perp$ di un qualunque sottoinsieme Y di S [pag.96] è un sottospazio chiuso di S [(1.9) pag.109], e che per ogni sottospazio H di S si ha $(H^\perp)^\perp = \overline{H}$ [(4.5) pag.120]. Ricordiamo altresì che se H è un sottospazio chiuso di S, lo spazio S è somma diretta di H ed $H^\perp$ [(4.4) pag.119], e perciò ciascuno dei sottospazi $H, H^\perp$ si chiama il supplemento ortogonale dell'altro. Ovviamente $S^\perp = \{0\}$ e $\{0\}^\perp = S$.

Si tenga presente che il nucleo $\operatorname{Ker} A$ di un operatore lineare e continuo $A: S\to S$ è un sottospazio chiuso di S, e che l'aggiunto dell'aggiunto di A coincide con A. Infine, si tenga presente che l'uguaglianza $\operatorname{Ker} A = \{0\}$ equivale al fatto che A è invertibile [pag.72-73].

Per comprendere bene il teorema che andiamo ad enunciare, è opportuno che il lettore si soffermi adeguatamente su quanto appena premesso.

(4.12) Si consideri un operatore lineare e continuo di S in S, con S spazio di Hilbert.

Il complemento ortogonale del codominio coincide con il nucleo dell'operatore aggiunto, e quindi, se il codominio è tutto lo spazio, l'operatore aggiunto è invertibile [in quanto il suo nucleo è $\{0\}$].



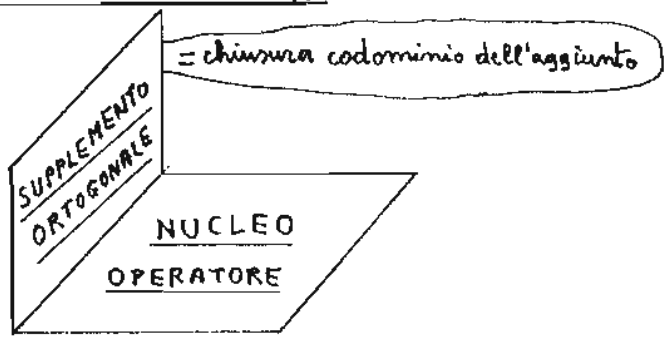
In simboli, denotando con B l'operatore in questione, si ha:

$$B(S)^{\perp} = \operatorname{Ker} B^{*} \tag{17}$$

e quindi:

$$B(S) = S \Longrightarrow B^{*} \text{ invertibile}.$$

Il complemento ortogonale del nucleo dell'operatore (supplemento ortogonale) coincide con la chiusura del codominio dell'aggiunto. Di conseguenza, se l'operatore è invertibile, il codominio dell'aggiunto è denso in S (sicché coincide con S se è chiuso).



In simboli, denotando con B l'operatore, si ha:

$$(\operatorname{Ker} B)^{\perp} = \overline{B^{*}(S)} \tag{18}$$

e quindi:

$$B \text{ invertibile} \Longrightarrow B^{*}(S) \text{ denso in } S.$$

DIM. Detto B l'operatore in questione, sussistono le equivalenze:

$$x \perp B(S) \Leftrightarrow (x,y)=0 \ \forall y \in B(S) \Leftrightarrow (x, Bx')=0 \ \forall x' \in S \Leftrightarrow$$
$$\Leftrightarrow (B^* x, x')=0 \ \forall x' \in S \Leftrightarrow B^* x = 0 \Leftrightarrow x \in \ker B^*$$

donde la (17).

Per dimostrare la seconda parte del teorema, applichiamo la prima parte all'operatore B^*. Si ha dunque:

$$B^*(S)^\perp = \ker B$$

e quindi

$$\left(B^*(S)^\perp \right)^\perp = (\ker B)^\perp$$

ossia:

$$\overline{B^*(S)} = (\ker B)^\perp$$

e la (18) è acquisita.

Dalla (4.12) si deduce il teorema:

(4.13). Si consideri un operatore lineare e continuo di S in S, con S spazio di Hilbert, e si supponga autoaggiunto.

Se il codominio è tutto lo spazio, l'operatore è invertibile.

Viceversa, se l'operatore è invertibile, il suo codominio è denso in S (sicché coincide con S se è chiuso).

DIM. Basta applicare la (4.12) tenendo presente che il complemento ortogonale di S è $\{0\}$, e il complemento ortogonale di $\{0\}$ è S; occorre ricordare che un operatore autoaggiunto è uguale al suo aggiunto.

ESEMPI SULLA NOZIONE DI OPERATORE AGGIUNTO

1) Un operatore lineare A di C^k in C^k si identifica con una matrice quadrata di ordine k, ed è lineare e continuo [quanto detto nell'es. 3 di pag. 64, e poi nell'es. 2 di pag. 145, si può ripetere sostituendo R con C].

Si verifica facilmente che l'aggiunto di A si identifica con la

trasposta della matrice coniugata di A. Denotando con a_{ij} l'elemento della riga i^{ma} e della colonna j^{ma} della matrice, se $a_{ij} = \bar{a}_{ji}$ $\forall i \in \{1,\dots,k\}$ e $\forall j \in \{1,\dots,k\}$ (sicché gli elementi della diagonale principale sono reali), risulta $A^* = A$ e quindi l'operatore è autoaggiunto.

Per verificare la precedente asserzione, supponiamo per semplicità $k=2$.

Qualunque siano $x=(x_1,x_2)\in C^2$, $y=(y_1,y_2)\in C^2$ si ha:

$$(19)\quad (x,Ay) = \big((x_1,x_2),(a_{11}y_1+a_{12}y_2,\ a_{21}y_1+a_{22}y_2)\big) =$$
$$= x_1(\bar{a}_{11}\bar{y}_1+\bar{a}_{12}\bar{y}_2)+x_2(\bar{a}_{21}\bar{y}_1+\bar{a}_{22}\bar{y}_2) = (\bar{a}_{11}x_1+\bar{a}_{21}x_2)\bar{y}_1+(\bar{a}_{12}x_1+\bar{a}_{22}x_2)\bar{y}_2$$
$$= \big((\bar{a}_{11}x_1+\bar{a}_{21}x_2,\ \bar{a}_{12}x_1+\bar{a}_{22}x_2),(y_1,y_2)\big).$$

Se si pone:

$$(20)\qquad A^* = \begin{pmatrix} \bar{a}_{11} & \bar{a}_{21} \\ \bar{a}_{12} & \bar{a}_{22} \end{pmatrix}$$

la (19) diviene:

$$(x,Ay) = (A^*x,y)$$

e pertanto l'aggiunto di A si identifica con la matrice A^*.

2) Detto Ω un aperto di R^k e fissato un multiindice α, consideriamo l'operatore lineare:

$$(21)\qquad A : x\in H_0^m(\Omega) \to D^\alpha x \in L^2(\Omega) \qquad (m\in N,\ |\alpha|\le m)$$

Evidentemente $H_0^m(\Omega)$ è denso in $L^2(\Omega)$ (1).

Detto y un arbitrario elemento di $H_0^m(\Omega)$, ricordando il significato di $D^\alpha y$ [(6) pag.61] si ha:

$$(22)\qquad \int_\Omega \varphi\, D^\alpha y\, dt = (-1)^{|\alpha|}\int_\Omega y\, D^\alpha\varphi\, dt \qquad \forall \varphi\in C_0^\infty(\Omega).$$

D'altra parte per ogni $x\in H_0^m(\Omega)$, essendo $C_0^\infty(\Omega)$ denso in $H_0^m(\Omega)$ [es.15 pag.103] esiste una successione $\{\varphi_n\}$ di punti di $C_0^\infty(\Omega)$ che tende ad x in $H_0^m(\Omega)$, sicché in $L^2(\Omega)$ si ha $\varphi_n\to x$, $D^\alpha\varphi_n\to D^\alpha x$. Dalla (22) scritta con $\varphi=\varphi_n$ consegue allora,

---

(1) Sia x_0 un arbitrario elemento di $L^2(\Omega)$ e sia I un qualunque intorno di x_0 [nello spazio $L^2(\Omega)$]. Essendo $C_0^\infty(\Omega)$ denso in $L^2(\Omega)$ [es.11 pag.102], all'interno I appartiene almeno un punto x di $C_0^\infty(\Omega)$; il punto x, identificato con una classe di funzioni, banalmente appartiene ad $H_0^m(\Omega)$.

Ciò prova che $H_0^m(\Omega)$ è denso in $L^2(\Omega)$.

passando al limite per $n\to\infty$:

$$(23)\qquad \int_\Omega x\, D^\alpha y\, dt = (-1)^{|\alpha|}\int_\Omega y\, D^\alpha x\, dt \qquad \forall x,y\in H_0^\infty(\Omega)$$

Denotando con $(\cdot,\cdot)$ il prodotto scalare in $L^2(\Omega)$, la (23) si può scrivere:

$$(x, D^\alpha y) = \left((-1)^{|\alpha|} D^\alpha x\, , y\right) \qquad \forall x,y\in H_0^\infty(\Omega).$$

Pertanto, se si considera l'operatore:

$$(24)\qquad A^*: x\in H_0^\infty(\Omega)\to (-1)^{|\alpha|} D^\alpha x\ ,$$

risulta

$$(x, Ay) = (A^* x, y) \qquad \forall x,y\in H_0^\infty(\Omega)$$

e quindi l'operatore (23) è l'aggiunto dell'operatore (21).

Se $|\alpha|$ è un numero pari risulta $A^*=A$ e quindi A è autoaggiunto.

Consideriamo ora l'operatore di Laplace in $H_0^2(\Omega)$:

$$\Delta_2 : x\in H_0^2(\Omega)\to \sum_{i=1}^{k} \frac{\partial^2 x}{\partial t_i^2} \in L^2(\Omega)$$

dove le derivate seconde vanno intese, ovviamente, in senso debole.

Trattandosi della somma di operatori del tipo (21) con $|\alpha|=2$, quindi di operatori autoaggiunti, l'operatore Δ_2 è autoaggiunto [cfr. (9) pag. 222].

3) Detto Ω un aperto di R^k, consideriamo l'operatore lineare:

$$(25)\qquad A: x\in H_0^2(\Omega)\to \sum_{i,j=1}^{1\dots k} \frac{\partial}{\partial t_i}\left(a_{ij}\frac{\partial x}{\partial t_j}\right) \in L^2(\Omega)$$

con i coefficienti a_{ij} di classe $C^1(\Omega)$.

Denotando con $(\cdot,\cdot)$ il prodotto scalare in $L^2(\Omega)$, per $x,y\in H_0^2(\Omega)$ si ha, tenuto conto della densità di $H_0^2(\Omega)$ in $L^2(\Omega)$:

$$(x,Ay)=\sum_{i,j=1}^{1\dots k}\int_\Omega x\frac{\partial}{\partial t_i}\left(a_{ij}\frac{\partial y}{\partial t_j}\right)dt = -\sum_{i,j=1}^{1\dots k}\int_\Omega a_{ij}\frac{\partial y}{\partial t_j}\frac{\partial x}{\partial t_i}\,dt = \sum_{i,j=1}^{1\dots k}\int_\Omega y\frac{\partial}{\partial t_j}\left(a_{ij}\frac{\partial x}{\partial t_i}\right)dt$$

ossia, scambiando nell'ultimo integrale il ruolo degli indici i,j:

$$(x,Ay)=\sum_{i,j}\int_\Omega y\frac{\partial}{\partial t_i}\left(a_{ji}\frac{\partial x}{\partial t_j}\right)dt\ .$$

Pertanto, se si introduce l'operatore:

$$(26)\qquad A^*: x\in H_0^2(\Omega)\to \sum_{i,j=1}^{1\dots k}\frac{\partial}{\partial t_i}\left(a_{ji}\frac{\partial x}{\partial t_j}\right)\in L^2(\Omega)$$

risulta

$$(x,Ay)=\left(A^*x,y\right) \qquad \forall x,y\in H_0^2(\Omega)$$

e quindi l'operatore (26) è l'aggiunto dell'operatore (25).

Se risulta $a_{ij}=a_{ji}$ qualunque siano $i,j\in\{1,\dots,k\}$, gli operatori A, A^* coincidono e quindi A è autoaggiunto.

Notiamo che l'operatore A è autoaggiunto anche nel caso che risulti:

$$\frac{\partial}{\partial t_i} a_{ij} = \frac{\partial}{\partial t_i} a_{ji} \quad \forall i,j \in \{1,\dots,k\}. \tag{27}$$

Per dimostrare tale asserzione scriviamo Ax nella forma seguente:

$$Ax = \frac{1}{2} \sum_{i,j}^{1\dots k} \left[\frac{\partial}{\partial t_i} \left(a_{ij} \frac{\partial x}{\partial t_j} \right) + \frac{\partial}{\partial t_j} \left(a_{ji} \frac{\partial x}{\partial t_i} \right) \right] =$$

$$= \frac{1}{2} \sum_{i,j}^{1\dots k} \left[(a_{ij} + a_{ji}) \frac{\partial^2 x}{\partial t_i \partial t_j} + \frac{\partial a_{ij}}{\partial t_i} \frac{\partial x}{\partial t_j} + \frac{\partial a_{ji}}{\partial t_j} \frac{\partial x}{\partial t_i} \right].$$

Siccome l'operatore (26) si ottiene formalmente sostituendo in (25) a_{ij} con a_{ji}, dall'espressione appena scritta sostituendo a_{ij} con a_{ji} (e quindi a_{ji} con a_{ij}) si ottiene $A^* x$. D'altra parte con tale sostituzione l'espressione appena scritta resta invariata se vale la (27), perché poi si può scrivere scambiando i ruoli degli indici i e j. Dunque risulta $A^* x = Ax \ \forall x \in H_0^1(\Omega)$ e pertanto $A = A^*$.

In particolare l'operatore A è autoaggiunto se i coefficienti a_{ij} sono costanti.

4) Consideriamo l'operatore lineare:

$$A : x \in H_0^2(\Omega) \to \sum_{i,j}^{1\dots k} a_{ij} \frac{\partial^2 x}{\partial t_i \partial t_j} \in L^2(\Omega)$$

con $a_{ij} \in C^2(\Omega)$, $\Omega \subseteq R^k$.

Denotando con $(\cdot,\cdot)$ il prodotto scalare in $L^2(\Omega)$, per $x, y \in H_0^2(\Omega)$ si ha:

$$(x, Ay) = \sum_{i,j}^{1\dots k} \int_\Omega x\, a_{ij} \frac{\partial^2 y}{\partial t_i \partial t_j}\, dt = -\sum_{i,j}^{1\dots k} \int_\Omega \frac{\partial y}{\partial t_i} \frac{\partial}{\partial t_j} (x a_{ij})\, dt = \sum_{i,j}^{1\dots k} \int_\Omega y \frac{\partial^2}{\partial t_i \partial t_j} (x a_{ij})\, dt$$

e quindi l'operatore che segue è l'aggiunto di A:

$$A^* : x \in H_0^2(\Omega) \to \sum_{i,j}^{1\dots k} \frac{\partial^2 (x a_{ij})}{\partial t_i \partial t_j} \in L^2(\Omega).$$

Notiamo che l'operatore A non è necessariamente autoaggiunto se $a_{ij} = a_{ji}$, mentre è autoaggiunto se i coefficienti a_{ij} sono tutti costanti. In quest'ultimo caso, peraltro, A coincide con l'operatore considerato nell'esempio precedente.

5) Consideriamo gli operatori integrali J_1 e J_2 generati da un nucleo $K(s,t)$ elemento di $L^2(\Omega\times\Omega)$ (pag. 168 e 335). Verifichiamo che:

L'aggiunto dell'operatore integrale J_1 [J_2] generato dal nucleo $K(s,t)$ con variabile di integrazione t [s] è l'operatore integrale generato dal nucleo $\overline{K(t,s)}$ con variabile di integrazione t [s].

Riferiamoci ad es. all'operatore J_1, che ad ogni $x\in L^2(\Omega)$ associa l'elemento v_x di $L^2(\Omega)$ definito ponendo:

$$v_x(s) = J_1x(s) = \int_\Omega K(s,t)x(t)\,dt \qquad \text{q.o. in } \Omega$$

Dobbiamo constatare che J_1^* è l'operatore che ad ogni $y\in L^2(\Omega)$ associa l'elemento w_y di $L^2(\Omega)$ definito ponendo:

$$w_y(s) = \int_\Omega \overline{K(t,s)}\,y(t)\,dt \qquad \text{q.o. in } \Omega. \tag{28}$$

Si ha, qualunque siano $x, y\in L^2(\Omega)$:

$$(J_1x, y) = \int_\Omega v_x(s)\overline{y}(s)\,ds = \int_\Omega \overline{y}(s)\,ds \int_\Omega K(s,t)x(t)\,dt = \iint_{\Omega\times\Omega} K(s,t)x(t)\overline{y}(s)\,ds\,dt =$$

$$= \int_\Omega x(t)\,dt \int_\Omega K(s,t)\overline{y}(s)\,ds = \int_\Omega x(t)\,dt \int_\Omega K(\xi,t)\overline{y}(\xi)\,d\xi .$$

D'altra parte la (28) si può scrivere anche

$$w_y(s) = \int_\Omega \overline{K(\xi,s)}\,y(\xi)\,d\xi$$

e quindi anche:

$$w_y(t) = \int_\Omega \overline{K(\xi,t)}\,y(\xi)\,d\xi .$$

Dunque il prodotto scalare dianzi calcolato si può scrivere:

$$(J_1x, y) = \int_\Omega x(t)\overline{w_y(t)}\,dt = (x, w_y)$$

e pertanto l'operatore che ad ogni $y\in L^2(\Omega)$ associa w_y è l'aggiunto di J_1.

Ovviamente se risulta:

$$K(s,t) = \overline{K(t,s)} \qquad \forall (s,t)\in\Omega\times\Omega \tag{29}$$

l'operatore J_1 è autoaggiunto.

Quando sussiste la (29) il nucleo K si dice <u>simmetrico</u>, oppure <u>hermitiano</u>.

5. Aggiunto di un operatore lineare e continuo fra spazi normati.

Nel numero precedente ci siamo occupati dell'aggiunto hermitiano di un operatore lineare, definito in un sottospazio denso di uno spazio di Hilbert. Come si è visto, abbiamo sistematicamente sottinteso l'attributo "hermitiano", conformemente al linguaggio corrente.

Peraltro si parla di operatore aggiunto, senza l'attributo "hermitiano", con riferimento ad un operatore lineare, che si suppone continuo, di uno spazio normato S in uno spazio normato S_1 (non necessariamente uguale ad S). Nel caso particolare $S=S_1$ spazio di Hilbert, l'aggiunto non si riduce all'aggiunto hermitiano, perchè come vedremo è un operatore di S^* in S^*, essendo S^* il duale di S; vedremo che coincide con l'aggiunto hermitiano nel caso che lo spazio di Hilbert S venga identificato con il suo duale (si ricordi il teor.(7.3) pag.180).

Allo scopo di introdurre la nozione accennata, denotiamo con S, S_1 due spazi vettoriali normati e consideriamo un operatore lineare

$$A : S \to S_1$$

che si suppone continuo.

Un elemento u_1 dello spazio S_1^* duale di S_1 (duale forte, pag.173) è un funzionale definito in S_1 (che ha coniugato lineare), e quindi per ogni $x \in S$ ha senso $\langle u_1, Ax \rangle$. Consideriamo allora il funzionale u definito in S ponendo:

$$u(x) = \langle u_1, Ax \rangle \quad \forall x \in S.$$

Il funzionale u ha ovviamente coniugato lineare, ed è continuo in quanto risulta:

$$(1) \qquad |u(x)| = |\langle u_1, Ax \rangle| \leq \|u_1\| \cdot \|Ax\| \leq \|u_1\| \cdot \|A\| \cdot \|x\| \quad \forall x \in S.$$

Pertanto u appartiene allo spazio S^* duale di S.

Col discorso precedente per ogni $u_1 \in S_1^*$ si viene ad individuare un elemento u di S^*, e perciò ha senso considerare la corrispondenza

$$A^* : u_1 \in S_1^* \to u \in S^*.$$

Dunque $(A^* u_1)(x) = u(x) = \langle u_1, Ax \rangle \ \forall x \in S$, cioè risulta

$$\langle A^* u_1, x \rangle = \langle u_1, Ax \rangle \qquad \forall x \in S \text{ e } \forall u_1 \in S_1^* . \tag{2}$$

Il lettore osservi che la (2) si può leggere come segue: per ogni $u_1 \in S_1^*$, $A^* u_1$ è il funzionale che in ogni punto $x \in S$ assume il valore $\langle u_1, Ax \rangle$; ossia, $A^* u_1$ è il funzionale che prima abbiamo indicato con u.

Si può dire perciò concisamente che A^* è la corrispondenza di S_1^* in S^* definita mediante l'uguaglianza (2).

Si verifica subito che A^* è un operatore lineare. Infatti qualunque siano $u_1, v_1 \in S_1^*$ si ha, tenendo presente la (2):

$$\langle A^*(u_1+v_1), x \rangle = \langle u_1+v_1, Ax \rangle = \langle u_1, Ax \rangle + \langle v_1, Ax \rangle =$$
$$= \langle A^* u_1, x \rangle + \langle A^* v_1, x \rangle = \langle A^* u_1 + A^* v_1, x \rangle \qquad \forall x \in S$$

e quindi $A^*(u_1+v_1) = A^* u_1 + A^* u_2$. Inoltre per ogni $\lambda \in C$ si ha:

$$\langle A^*(\lambda u_1), x \rangle = \langle \lambda u_1, Ax \rangle = \lambda \langle u_1, Ax \rangle = \lambda \langle A^* u_1, x \rangle$$
$$= \langle \lambda A^* u_1, x \rangle \qquad \forall x \in S$$

e quindi $A^*(\lambda u_1) = \lambda A^* u_1$.

L'operatore A^* è anche continuo. Infatti dalla (1), il primo membro della quale è $A^* u_1$, si deduce:

$$\frac{|(A^* u_1)(x)|}{\|x\|} \leq \|A\| \cdot \|u_1\|$$

e quindi l'estremo superiore del primo membro per $x \in S - \{0\}$ è $\leq \|A\| \cdot \|u_1\|$, cioè risulta [(2) pag. 202]:

$$\|A^* u_1\| \leq \|A\| \cdot \|u_1\| \qquad \forall u_1 \in S_1^*$$

donde la continuità dell'operatore A^*.

Orbene l'operatore lineare e continuo $A^*: S_1^* \to S^*$ definito mediante la (2) si chiama l'operatore aggiunto di A. Dalla precedente uguaglianza consegue che $\|A^*\| \leq \|A\|$, ma è facile verificare che risulta:

$$\|A^*\| = \|A\| . \tag{3}$$

DIM. Detto x un punto di S tale che $Ax \neq 0$, per un corollario del teorema di Hahn-Banach [I (9.3) pag. 184] esiste $u_1 \in S_1^* - \{0\}$ tale da aversi $|\langle u_1, Ax \rangle| = \|u_1\| \cdot \|Ax\|$. Si ha allora:

$$\|Ax\| = \frac{|\langle u_1, Ax \rangle|}{\|u_1\|} = \frac{|\langle A^* u_1, x \rangle|}{\|u_1\|} \leq \frac{\|A^* u_1\| \cdot \|x\|}{\|u_1\|} \leq \|A^*\| \cdot \|x\|$$

da cui segue $\|A\| \leq \|A^*\|$, donde la (3).

Ovviamente, detto S uno spazio vettoriale normato, l'aggiunto dell'identità in S è l'identità in S^*.

In base alla definizione si verifica subito che:

(5.1). Se A, B sono operatori lineari e continui, e $\lambda \in K$, si ha:

$$(4) \qquad (A+B)^* = A^* + B^* \quad , \quad (AB)^* = B^* A^* \quad , \quad (\lambda A)^* = \bar{\lambda} A^* .$$

Dalla seconda delle (4) consegue poi:

(5.2). Se A è un isomorfismo di S su S_1, anche il suo aggiunto è un isomorfismo (di S_1^* su S^*); e l'inverso dell'aggiunto di A è uguale all'aggiunto dell'inverso di A.

Detta I l'identità in S ed I_1 l'identità in S_1, dalle uguaglianze

$$I = A^{-1}A \quad , \quad I_1 = AA^{-1}$$

consegue:

$$I^* = A^*(A^{-1})^* \quad , \quad I_1^* = (A^{-1})^* A^* .$$

Dunque A^* è dotato di inverso sinistro e di inverso destro, e pertanto A^* è invertibile ed ha per codominio tutto lo spazio. Di conseguenza l'inverso di A^*, definito in tutto S^*, è uguale all'inverso destro e all'inverso sinistro, ossia ad $(A^{-1})^*$.

Dalla (5.2) si trae che se due spazi vettoriali normati sono isomorfi, anche i rispettivi spazi duali sono isomorfi.

Nei teoremi seguenti, che ci limitiamo ad enunciare, denotiamo con A un operatore lineare e continuo di S in S_1, essendo S, S_1 spazi di Banach:

(5.3) Sussistono le equivalenze:

$$A \text{ compatto} \iff A^* \text{ compatto} \qquad (1)$$

$$\left(\begin{array}{c}\text{codominio di } A \\ \text{di dimensione finita } n\end{array}\right) \iff \left(\begin{array}{c}\text{codominio di } A^* \\ \text{di dimensione finita } n\end{array}\right) \qquad (2)$$

e nell'ipotesi che $A : S \to S$ sia compatto, per ogni $\lambda \neq 0$:

$$\ker(A - \lambda I) \text{ ha dimensione } n \iff \ker(A^* - \bar{\lambda} I^*) \text{ ha dimensione } n \qquad (3)$$

---

(1) L'implicazione $\Longrightarrow$ sussiste anche se S, S_1 non sono completi, quella opposta anche se S non è completo.
(2) L'equivalenza sussiste anche se S, S_1 non sono completi.
(3) Conveniamo che il sottospazio $\{0\}$ abbia dimensione nulla. Per il teor. (5.2) di pag. 166 i due nuclei hanno dimensione finita (eventualmente nulla). La dichiarata equivalenza sussiste anche per $n = 0$, e quindi comprende la seguente: $A - \lambda I$ invertibile $\Leftrightarrow$ $A^* - \bar{\lambda} I^*$ invertibile.

Sussiste poi il teorema:

(5.4). Qualunque sia B operatore lineare e continuo di S in S_1, con S ed S_1 spazi vettoriali normati, si ha quanto segue.

Il nucleo di B [di B^*] coincide con l'insieme degli elementi ortogonali al codominio di B^* [di B].

Il codominio di B [di B^*] è contenuto nell'insieme degli elementi ortogonali al nucleo di B^* [di B], e coincide con tale insieme se il codominio di B è chiuso [se il codominio di B^* è chiuso (1)]

Dimostriamo la prima affermazione dell'enunciato.

Per $x \in S$ si ha:

x ortogonale al codominio di $B^* \iff \langle u, x\rangle = 0 \ \forall u \in B^*(S_1^*) \iff \langle B^* u_1, x\rangle = 0 \ \forall u_1 \in S_1^* \iff$

$\iff \langle u_1, Bx\rangle = 0 \ \forall u_1 \in S_1^* \iff Bx = 0$ (2) $\iff x \in \ker B$.

Per $u_1 \in S_1^*$ si ha:

u_1 ortogonale al codominio di $B \iff \langle u_1, y\rangle = 0 \ \forall y \in B(S) \iff \langle u_1, Bx\rangle = 0 \ \forall x \in S \iff$

$\iff \langle B^* u_1, x\rangle = 0 \ \forall x \in S \iff B^* u_1 = 0 \iff u_1 \in \ker B^*$.

Della seconda affermazione dell'enunciato, dimostriamo la prima parte.

Osserviamo preliminarmente che un elemento y di S_1 è ortogonale al nucleo di B^* quando per ogni $u_1 \in \ker B^*$ risulta $\langle u_1, y\rangle = 0$, cioè quando sussiste l'implicazione:

(5) $$B^* u_1 = 0 \Rightarrow \langle u_1, y\rangle = 0 .$$

Si tratta allora di far vedere che se y appartiene al codominio di B sussiste la (5). Infatti, se $y \in B(S)$ esiste $x \in S$ tale che $y = Bx$, e quindi:

$$B^* u_1 = 0 \Rightarrow \langle B^* u_1, x\rangle = 0 \iff \langle u_1, Bx\rangle = 0 \iff \langle u_1, y\rangle = 0 .$$

Analogamente, un elemento u di S^* è ortogonale al nucleo di B quando per ogni $x \in \ker B$ risulta $\langle u, x\rangle = 0$, cioè quando sussiste l'implicazione:

(6) $$Bx = 0 \Rightarrow \langle u, x\rangle = 0 .$$

Si tratta allora di far vedere che se u appartiene al codominio di B^* sussiste la (6). Infatti, se $u \in B^*(S_1^*)$ esiste $u_1 \in S_1^*$ tale che $u = B^* u_1$, e quindi:

$$Bx = 0 \Rightarrow \langle u_1, Bx\rangle = 0 \iff \langle B^* u_1, x\rangle = 0 \iff \langle u, x\rangle = 0 .$$

Dimostriamo ora che, se $B(S)$ è chiuso, ogni elemento y di S_1 ortogonale a $\ker B^*$, ossia tale da verificare la (5), appartiene a $B(S)$. Procedendo per assurdo, ammettiamo che esista un elemento y di S_1 per il quale sussiste la (5), che non appartiene al codominio di B, diciamolo H.

Siccome H è chiuso ed esiste $y \notin H$, il sottospazio H di S_1 non è denso in S_1 ed allora è applicabile il corollario (9.2) di pag. 183 del teorema di Hahn-Banach: esiste dunque $u_1 \in S_1^*$ tale che:

(7) $$\langle u_1, y\rangle \neq 0$$

(7') $$\langle u_1, h\rangle = 0 \quad \forall h \in H .$$

La (7') equivale a $\langle u_1, Bx\rangle = 0 \ \forall x \in S$, ossia $\langle B^* u_1, x\rangle = 0 \ \forall x \in S$, cioè:

(1) L'asserto nel caso che il codominio di B^* è chiuso sarà dimostrato, per semplicità, nell'ulteriore ipotesi che S sia uno spazio di Banach riflessivo. Segnaliamo, senza dimostrazione, che se S, S_1 sono spazi di Banach, per ogni $B \in \mathcal{L}(S, S_1)$ si ha: B a codominio chiuso $\iff$ B^* a codominio chiuso.

(2) Per un corollario del teorema di Hahn-Banach [cfr. (9.4) pag. 184 con $x_2 = 0$], se in un punto y di uno spazio normato si annullano tutti gli elementi del duale, risulta necessariamente $y = 0$.

(8) $\qquad B^* u_1 = 0$

Le (7), (8) contraddicono l'ipotesi (5), e l'asserto è provato.

Resta da far vedere che, supposto S riflessivo e $B^*(S_1^*)$ chiuso, ogni elemento u di S^* ortogonale al nucleo di B, ossia tale che risulti la (6), appartiene a $B^*(S_1^*)$. Per assurdo, ammettiamo che esista $u \in S^*$ per il quale valga la (6), che non appartiene al codominio di B^*, diciamolo H. Siccome H è chiuso e non coincide con S^*, il sottospazio H di S^* non è denso in S^* e quindi esiste, per un corollario del teor. di Hahn-Banach [(9.2) pag. 183], un elemento v di S^{**} tale che $\langle\langle v, u \rangle\rangle \neq 0$ e $\langle\langle v, h \rangle\rangle = 0 \ \forall h \in H$. Essendo S riflessivo, esiste allora un elemento x di S tale che:

(8') $\qquad \langle u, x \rangle \neq 0 \quad , \quad \langle h, x \rangle = 0 \ \forall h \in H.$

La seconda delle (8') significa che $\langle B^* u_1, x \rangle = 0 \ \forall u_1 \in S_1^*$, ossia $\langle u_1, Bx \rangle = 0 \ \forall u_1 \in$: ciò implica che $Bx = 0$ [nota (2) pag. 230] e quindi, per la (6), che $\langle u, x \rangle = 0$. Quest'ultima uguaglianza è assurda per la prima delle (8'), e l'asserto è dimostrato.

Un'importante conseguenza del teorema (5.4) è espressa dal seguente:

(5.5). Sia B un operatore lineare e continuo di S in S_1, con S, S_1 spazi vettoriali normati.

Se B [B^*] ha per codominio tutto lo spazio, allora B^* [B] è invertibile.

Viceversa, nell'ipotesi che il codominio di B [di B^*] sia chiuso, se B^* [B] è invertibile l'operatore B [B^*] ha per codominio tutto lo spazio.

DIM. Per la (5.4) si ha:

$$\operatorname{Ker} B^* = \{u_1 \in S_1^* : u_1 \perp B(S)\} \quad , \quad \operatorname{Ker} B = \{x \in S : x \perp B^*(S_1^*)\}.$$

Se $B(S) = S$ si ha allora:

$$u_1 \in \operatorname{Ker} B^* \Leftrightarrow \langle u_1, y \rangle = 0 \ \forall y \in S_1 \Leftrightarrow u_1 = 0$$

e quindi $\operatorname{Ker} B^* = \{0\}$, ossia B^* è invertibile. Analogamente, se $B^*(S_1^*) = S^*$, si ha:

$$x \in \operatorname{Ker} B \Leftrightarrow \langle u, x \rangle = 0 \ \forall u \in S^* \Leftrightarrow u = 0 \ \text{[nota (2) pag. 230]}$$

e quindi $\operatorname{Ker} B = \{0\}$, ossia B è invertibile.

Per dimostrare il viceversa, supponiamo $B(S)$ chiuso [$B^*(S_1^*)$ chiuso]. Si ha per la (5.4):

$$B(S) = \{y \in S_1 : y \perp \mathrm{Ker}\, B^*\}, \quad B^*(S_1^*) = \{u \in S^* : u \perp \mathrm{Ker}\, B\}.$$

Se $\mathrm{Ker}\, B^* = \{0\}$ si ha allora:

$$y \in B(S) \Leftrightarrow y \perp 0 \Leftrightarrow y \in S_1$$

e quindi $B(S) = S_1$. Analogamente se $\mathrm{Ker}\, B = \{0\}$, si ha:

$$u \in B^*(S_1^*) \Leftrightarrow u \perp 0 \Leftrightarrow u \in S^*$$

e quindi $B^*(S_1^*) = S^*$.

Supponiamo ora che S, S_1 siano spazi di Hilbert, e denotiamo con j [con j_1] l'isomorfismo canonico di S^* su S [di S_1^* su S_1].

Se A è un operatore lineare e continuo di S in S_1, la relazione:

$$\langle u_1, Ax \rangle = \langle A^* u_1, x \rangle \quad \forall u_1 \in S_1^* \text{ e } \forall x \in S$$

ponendo $j_1 u_1 = y$ si scrive:

$$(y, Ax) = (jA^* j_1^{-1} y, x) \quad \forall y \in S_1 \text{ e } \forall x \in S.$$

Se allora si pone $A' = jA^* j_1^{-1}$, l'operatore $A' : S_1 \to S$ verifica la relazione:

$$(9) \qquad (y, Ax) = (A'y, x) \quad \forall y \in S_1 \text{ e } \forall x \in S.$$

Nel caso particolare $S_1 = S$ l'operatore A' soddisfacente alla (9) è l'aggiunto hermitiano di A: perciò in ogni caso A' si chiama l'aggiunto hermitiano di $A : S \to S_1$, e come già evidenziato è un operatore di S_1 in S.

Tornando al caso particolare $S_1 = S$, abbiamo già detto che spesso l'attributo "hermitiano" si sottintende: perciò è bene sottolineare che l'aggiunto hermitiano è diverso dall'aggiunto, dato che quest'ultimo è un operatore di S^* in S^*; nel numero precedente avremmo dovuto adoperare il simbolo A' invece di A^*, ma quest'ultimo è il simbolo abitualmente usato. Notiamo che se lo spazio S si identifica con S^* [cfr. (7.3) pag. 180], l'operatore $j = j_1$ si identifica con l'operatore identico e quindi $A' = jA^* j^{-1}$ diviene $A' = A^*$.

6. Il teorema dell'alternativa.

Siano: S uno spazio di Banach complesso, A un operatore lineare e continuo di S in S, λ un numero complesso non nullo.

Fissato $x' \in S$, consideriamo l'equazione nell'incognita $x \in S$:

$$Ax - \lambda x = x' \tag{1}$$

il cui primo membro è il valore che assume nel punto x l'operatore $A-\lambda I$, essendo I l'identità in S.

Detta I^* l'identità nello spazio S^* duale di S, l'operatore aggiunto di $A-\lambda I$ è $A^*-\bar{\lambda} I^*$ [cfr. (5.1) pag. 229], ed è un operatore di S^* in S^*; perciò, fissato un elemento u' di S^*, l'equazione nell'incognita $u \in S^*$:

$$A^*u - \bar{\lambda} u = u' \tag{2}$$

si chiama <u>l'equazione aggiunta</u> della (1).

Le equazioni con "termine noto" nullo (equazioni <u>omogenee</u>):

$$Ax - \lambda x = 0 \tag{3}$$

$$A^*u - \bar{\lambda} u = 0 \tag{4}$$

si chiamano le <u>equazioni omogenee associate</u>, rispettivamente <u>alla (1)</u> e <u>alla (2)</u>.

Sussiste il seguente teorema di <u>Riesz-Fredholm</u>, noto come <u>teorema dell'alternativa</u>:

(6.1) <u>Con riferimento all'equazione (1) e all'equazione aggiunta (2), se l'operatore A è compatto si verifica necessariamente una delle due seguenti eventualità:</u>

1) <u>Le equazioni omogenee associate (3) e (4) ammettono entrambe soltanto la soluzione nulla.</u>

<u>Allora le equazioni (1) e (2) ammettono entrambe, qualunque siano i termini noti, una e una sola soluzione.</u>

2) <u>Le equazioni omogenee associate (3) e (4) ammettono infinite soluzioni, ed esiste un $n \in \mathbb{N}$ che è per entrambe il massimo numero di soluzioni linearmente indipendenti.</u>

<u>Allora l'equazione (1) [l'equazione aggiunta (2)] ammette soluzioni se e solo se il termine noto è ortogonale alle soluzioni dell'omogenea associata all'equazione aggiunta (2) [all'equazione (1)].</u>

Pertanto nell'eventualità 2), dette $x_1,\dots,x_n$ ed $u_1,\dots,u_n$ rispettivamente n soluzioni linearmente indipendenti delle (3) e (4), l'equazione (1) [l'equazione (2)] ammette soluzioni se e solo se il termine noto x' [il termine noto u'] soddisfa alle seguenti condizioni (condizioni "di compatibilità"):

$$(5) \qquad \langle u_i, x' \rangle = 0 \qquad [\langle u', x_i \rangle = 0] \qquad (i=1,\dots,n)$$

Quando tali condizioni sono verificate, le soluzioni della (1) [della (3)] sono date dalla formula:

$$(6) \qquad x = \tilde{x} + \sum_{i=1}^{n} c_i x_i \qquad \left[u = \tilde{u} + \sum_{i=1}^{n} c_i u_i \right]$$

dove $\tilde{x}$ [$\tilde{u}$] è una soluzione della stessa (1) [(2)], e $c_1,\dots,c_n$ sono costanti (complesse) arbitrarie.

— — — — — — — — — — — —

Dim. Le soluzioni della (3) [della (4)] sono gli elementi del nucleo dell'operatore $A-\lambda I$ [$A^*-\bar{\lambda} I^*$]. Essendo A compatto per ipotesi ed A^* compatto per la (5.3) di pag. 229, i due nuclei sono entrambi a dimensione finita [(5.2) pag. 166], e per la predetta (5.3) hanno la stessa dimensione: è chiaro allora che le eventualità 1) e 2) sono le sole possibili.

Nell'eventualità 1) l'operatore $B = A-\lambda I$ è invertibile, e l'operatore aggiunto $B^* = A^*-\bar{\lambda} I^*$ ha il codominio chiuso per il teorema (5.2) di pag. 166, sicché B^* ha per codominio tutto lo spazio S^*, in virtù del teor. (5.5) di pag. 231: ciò vuol dire che l'equazione aggiunta (2) ammette soluzioni qualunque sia il termine noto. D'altra parte se $u^{(1)}$ e $u^{(2)}$ sono soluzioni della (2), da $B^* u^{(1)} = u'$, $B^* u^{(2)} = u'$ segue $B^*[u^{(1)} - u^{(2)}] = 0$ e quindi, per la supposta eventualità 1), $u^{(1)} = u^{(2)}$. Pertanto l'equazione aggiunta (2) ammette, qualunque sia il termine noto, una ed una sola soluzione. Il ragionamento ora concluso si può ripetere scambiando i ruoli degli operatori B e B^*, tenuto conto della prima equivalenza del teor. (5.3) di pag. 229: se ne deduce che l'equazione (1) ammette una ed una sola soluzione, qualunque sia il termine noto.

Supponiamo ora che si verifichi l'eventualità 2).

Essendo A, e quindi A^*, operatori compatti, per il teor. (5.2) di pag. 166 gli operatori $B = A-\lambda I$ e $B^* = A^*-\bar{\lambda} I^*$ hanno entrambi il codominio chiuso. Ne consegue, per il teorema (5.4) di pag. 230, che un elemento x' di S [un elemento u' di S^*] appartiene al codominio di B [di B^*] se e solo se è ortogonale al nucleo di B^* [di B]: ciò equivale all'enunciato, perché l'appartenenza di x' [di u'] al codominio di B [di B^*] significa che l'equazione (1) [(2)] ammette soluzioni, e gli elementi del nucleo di B^* [di B] sono le soluzioni della (4) [(3)].

La conseguenza del teorema, formulata dopo l'enunciato, è pressocché ovvia. Siccome $(x_1,\dots,x_n)$ è una base di $\ker B$ [$(u_1,\dots,u_n)$ è una base di $\ker B^*$], il fatto che u' sia ortogonale agli x_i [che x' sia ortogonale agli u_i] equivale al fatto che u' [x'] è ortogonale a $\ker B$ [a $\ker B^*$], e naturalmente gli elementi di $\ker B$ [di $\ker B^*$] sono le combinazioni lineari degli x_i [degli u_i].

Se $B\tilde{x} = x'$, per ogni x tale che $Bx = x'$ si ha $B(x-\tilde{x}) = 0$, sicché $x-\tilde{x}$ è soluzione dell'equazione omogenea associata; viceversa se $B(x-\tilde{x}) = 0$, essendo $B\tilde{x} = x'$ anche $Bx = x'$. Analogo discorso si può fare per B^*, e dopo ciò è chiaro che la (6) rappresenta le soluzioni della (1) [della (2)].

— — — — — — — — — — — —

Se S è uno spazio di Hilbert, l'aggiunto (hermitiano) di A è un operatore di S in S (anzicchè di S^* in S^*), e quindi l'equazione aggiunta della (1):

$$A^*y - \bar{\lambda} y = y' \tag{7}$$

ha termine noto appartenente ad S; le soluzioni vanno ricercate in S (come per la (1)). Il teorema dell'alternativa sussiste anche se A^* è l'aggiunto hermitiano di A, sostituendo ivi la (2) con la (7) e la (4) con l'equazione omogenea associata alla (7):

$$A^*y - \bar{\lambda} y = 0 \,. \tag{8}$$

Naturalmente le (5) divengono

$$(y_i, x') = 0 \qquad [(y', x_i) = 0] \qquad (i = 1, 2, \ldots, n)$$

essendo $y_1, \ldots y_n$ n soluzioni linearmente indipendenti della (8) e $(\cdot,\cdot)$ il prodotto scalare in S.

Per la dimostrazione del teorema dell'alternativa nel caso S spazio di Hilbert vale quella esposta nel caso generale (con le varianti appena indicate): in luogo dei teoremi (5.4), (5.5) si utilizzerà il teorema (4.12) pag. 226; il teorema (5.3), anch'esso utilizzato nella dimostrazione, sussiste pure se, supponendo S, S_1 spazi di Hilbert, per operatore aggiunto si intende l'aggiunto hermitiano.

Sempre nel caso S spazio di Hilbert, supponiamo ora che l'operatore A sia autoaggiunto, cioè che sia $A^* = A$. L'equazione (8) può essere allora sostituita con l'equazione $Ay - \lambda y = 0$, che ha le stesse soluzioni della (8) con $A^* = A$ (1). Pertanto con riferimento ad un operatore autoaggiunto in uno spazio di Hilbert il teorema dell'alternativa si enuncia come segue.

(6.2) Sia A un operatore lineare e continuo di S spazio di Hilbert

(1) Ciò è banale se λ è reale. Se λ non è reale, le due equazioni hanno entrambe l'unica soluzione $y = 0$. Infatti essendo A autoaggiunto, per ogni $y \in S$ il numero (Ay, y) è reale [14.7) pag. 222 bis, con H = S]: se risulta $Ay - \lambda y = 0$, essendo $(Ay, y) = \lambda \|y\|^2$ si ha necessariamente $y = 0$ (per $y \neq 0$ si avrebbe (Ay, y) non reale, in quanto λ non è reale); il ragionamento vale anche se risulta $Ay - \bar{\lambda} y = 0$.

in S, e sia λ un numero complesso non nullo. Se A è compatto ed autoaggiunto, per l'equazione nell'incognita $x \in S$:

(9) $$Ax - \lambda x = 0$$

si verifica necessariamente una delle seguenti eventualità:

1) La (9) ammette soltanto la soluzione $x = 0$.

2) La (9) ammette infinite soluzioni, ed esiste un $n \in \mathbb{N}$ che è il massimo numero di soluzioni linearmente indipendenti. Detta $(x_1, \dots, x_n)$ una n-pla di siffatte soluzioni, le soluzioni della (9) sono date dalla formula $x = c_1 x_1 + c_2 x_2 + \dots + c_n x_n$, con $c_1, \dots, c_n$ costanti arbitrarie.

Nell'eventualità 1), qualunque sia $x' \in S$ l'equazione nell'incognita $x \in S$:

(10) $$Ax - \lambda x = x'$$

ammette una e una sola soluzione.

Nell'eventualità 2) l'equazione (10) ammette soluzioni se e solo se risulta:

(11) $$(x', x_i) = 0 \quad \forall i \in \{1, \dots, n\}.$$

Quando le (11) (dette condizioni di compatibilità) sono soddisfatte, le soluzioni della (10) sono date dalla formula:

(12) $$x = c_1 x_1 + \dots + c_n x_n + x^*$$

dove x^* è una soluzione della stessa (10) e $c_1, \dots, c_n$ sono costanti arbitrarie.

Esempio.

Nell'es. 1 di pag. 168 abbiamo introdotto l'equazione integrale di Fredholm:

(13) $$\int_\Omega k(\sigma, t) x(t)\, dt - \lambda x(\sigma) = \varphi(\sigma)$$

della quale ci occuperemo nell'Appendice [pag. 335]. Vogliamo qui osservare che con riferimento alla (13), e all'equazione aggiunta [es. 5 pag. 227.4]:

(14) $$\int_\Omega \overline{k(t, \sigma)}\, y(t)\, dt - \lambda y(\sigma) = \psi(\sigma)$$

sussiste la tesi del teorema dell'alternativa con S spazio di Hilbert [cfr. (6.1) pag. 233 con le varianti indicate a pag. 235].

Due sono le possibili eventualità. Nella prima, le (13), (14) ammettono entrambe, qualunque siano i termini noti, una e una sola soluzione.

Nella seconda, se $x_1, \dots, x_n$ [$y_1, \dots, y_n$] sono n soluzioni linearmente indipendenti

dell'omogenea associata alla (13) [alla (14)], l'equazione (13) [(14)] ammette soluzioni se e solo se il termine noto φ [ψ] verifica le uguaglianze:

$$(15) \qquad \int_\Omega \varphi(s)\,\overline{y_i(s)}\,ds = 0 \quad \left[\int_\Omega \psi(t)\,\overline{x_i(t)}\,dt = 0\right] \qquad (i=1,\dots,n).$$

Soddisfatte tali condizioni, le soluzioni x della (13) [y della (14)] sono date dalla formula:

$$(16) \qquad x = \tilde{x} + \sum_{i=1}^{n} c_i x_i \qquad \left[y = \tilde{y} + \sum_{i=1}^{n} c_i y_i\right]$$

dove $\tilde{x}$ [$\tilde{y}$] è una soluzione della stessa (13) [(14)] e $c_1 \dots c_n$ sono costanti arbitrarie.

Abbiamo visto a pag. 218 che, se $|\lambda|$ è maggiore della norma in $L^2(\Omega\times\Omega)$ del nucleo $K(s,t)$, si verifica la prima delle precedenti eventualità e la soluzione della (13) o della (14) si può esprimere mediante un operatore integrale, generato da un opportuno nucleo $H_\lambda(s,t)$.

Esercizio

Assegnata una funzione φ di quadrato sommabile nell'intervallo $(0,1)$, consideriamo per $\lambda \neq 0$ l'equazione integrale:

$$(17) \qquad \int_0^1 (s+t)\,x(t)\,dt - \lambda x(s) = \varphi(s).$$

L'operatore integrale J generato dal nucleo $K(s,t) = s+t$ è autoaggiunto [cfr. (24) pag. 227], ed associa ad ogni $x \in L^2[(0,1)]$ la funzione:

$$y(s) = s\int_0^1 x(t)\,dt + \int_0^1 t\,x(t)\,dt \quad ,$$

sicché il codominio di J è costituito da funzioni del tipo $as+b$ con a,b costanti.

Consideriamo l'equazione omogenea associata alla (17):

$$(18) \qquad \int_0^1 (s+t)\,x(t)\,dt - \lambda x(s) = 0$$

ed osserviamo che le soluzioni della (18), dovendo verificare l'uguaglianza $x = \frac{1}{\lambda}J(x)$, debbono appartenere al codominio di J e quindi debbono essere del tipo $as+b$. Poniamo allora nella (18) $x(t) = at+b$, al fine di stabilire per quali valori di a e b tale funzione verifica la (18). Si trova, con semplici calcoli:

$$\left(\frac{a}{2}+b-\lambda a\right)s + \left[\frac{a}{3}+\frac{b}{2}-\lambda b\right] = 0 \qquad \forall s \in [0,1]$$

e questa relazione è verificata se e solo se risulta:

$$(19) \qquad \left(\frac{1}{2}-\lambda\right)a + b = 0 \quad , \quad \frac{1}{3}a + \left(\frac{1}{2}-\lambda\right)b = 0 \quad .$$

Pertanto, se il determinante

$$(20) \qquad \begin{vmatrix} \frac{1}{2}-\lambda & 1 \\ \frac{1}{3} & \frac{1}{2}-\lambda \end{vmatrix}$$

è diverso da 0, il sistema (19) ammette l'unica soluzione $a=b=0$, e quindi la

(18) ammette soltanto la soluzione nulla. Dunque se λ è diverso dai valori:

$$\lambda_1 = \frac{1}{2} + \frac{\sqrt{3}}{3} \quad , \quad \lambda_2 = \frac{1}{2} - \frac{\sqrt{3}}{3} \tag{21}$$

l'equazione integrale (17) ammette, qualunque sia il termine noto φ, una e una sola soluzione.

Per determinare tale soluzione, poichè la (17) può scriversi $Jx - \lambda x = \varphi$, ossia $x = -\frac{1}{\lambda}\varphi + \frac{1}{\lambda} Jx$, ricordando che il codominio di J è costituito da polinomi al più di primo grado, poniamo:

$$x(t) = -\frac{1}{\lambda}\varphi(t) + \alpha t + \beta \tag{22}$$

e determiniamo α e β in modo che sia soddisfatta la (17). Sostituendo la (22) in (17) si ottiene il seguente sistema nelle incognite α e β:

$$\left(\frac{1}{2} - \lambda\right)\alpha + \frac{1}{3}\beta = \frac{1}{\lambda}\int_0^1 t\varphi(t)\,dt, \qquad \alpha + \left(\frac{1}{2} - \lambda\right)\beta = \frac{1}{\lambda}\int_0^1 \varphi(t)\,dt \tag{23}$$

che ammette soluzione unica in quanto abbiamo supposto il determinante (20) diverso da 0. In definitiva per ogni λ diverso dai valori (21) la (17) ammette l'unica soluzione (22) con (α, β) soluzione del sistema (23).

Per $\lambda = \frac{1}{2} + \frac{\sqrt{3}}{3}$ l'equazione (18) ammette infinite soluzioni, che sono del tipo [1] $x(s) = h(1 + s\sqrt{3})$ con h costante, sicchè il massimo numero di soluzioni linearmente indipendenti è . Dunque per il suddetto valore di λ l'equazione (17) ammette soluzioni se e solo se il termine noto φ soddisfa alla condizione:

$$\int_0^1 (1 + \sqrt{3}\,t)\varphi(t)\,dt = 0 . \tag{24}$$

Nell'ipotesi (24), risolvendo il sistema (23) nelle incognite α e β e sostituendo nella (22) si trovano le soluzioni della (17):

$$x(s) = h(1 + s\sqrt{3}) - \frac{2\sqrt{3}}{2 + \sqrt{3}}\left[\varphi(s) - \int_0^1 \varphi(t)\,dt\right]$$

con h costante arbitraria.

Analogamente, per $\lambda = \frac{1}{2} - \frac{\sqrt{3}}{3}$ la (17) ammette soluzioni se risulta:

$$\int_0^1 (1 - \sqrt{3}\,t)\varphi(t)\,dt = 0 ,$$

e in tale ipotesi le soluzioni sono date dalla formula:

$$x(s) = h(s\sqrt{3} - 1) + \frac{2\sqrt{3}}{2 - \sqrt{3}}\left[\varphi(s) - \int_0^1 \varphi(t)\,dt\right]$$

con h costante arbitraria.

---

[1] Come si è detto, sono del tipo $as + b$ con (a, b) soluzione del sistema (19): perciò si ottengono risolvendo il sistema (19) con $\lambda = \frac{1}{2} + \frac{\sqrt{3}}{3}$.

Capitolo VIII

Cenni sulla teoria spettrale.

1. Insieme risolvente e spettro di un operatore lineare di uno spazio vettoriale topologico in sé. Autovalori, autovettori.

Siano: S uno spazio vettoriale topologico sul campo complesso, A un operatore lineare di S in S (non necessariamente continuo), I l'identità in S. Per ogni $\lambda \in C$ poniamo:

$$A_\lambda = A - \lambda I .$$

Si chiama <u>insieme risolvente</u> dell'operatore A, e si denota con $\rho(A)$, il sottoinsieme di C costituito dai numeri complessi λ tali da verificare la seguente condizione:

<u>A_λ è dotato di inverso continuo, definito in un sottospazio di S denso in S</u>

ossia tali che si abbia:

(1)
a) <u>A_λ invertibile</u>.
b) <u>Codominio di A_λ denso in S</u>
c) <u>Inverso di A_λ continuo.</u>

L'insieme complementare di $\rho(A)$, cioè l'insieme costituito dai numeri complessi λ che non appartengono all'insieme risolvente di A, si chiama <u>lo spettro</u> dell'operatore A, e si denota con $\sigma(A)$.

Lo spettro di A si decompone in tre parti: spettro puntuale, spettro continuo, spettro residuo, che si denotano rispettivamente con $\sigma_p(A)$, $\sigma_c(A)$, $\sigma_r(A)$. Per definire tali parti di $\sigma(A)$, premettiamo il seguente schema.

EVENTUALITÀ POSSIBILI			PER DEFINIZIONE λ APPARTIENE A	SIMBOLO
A_λ invertibile	Codominio di A_λ denso in S	Inverso di A_λ continuo	insieme risolvente	$\rho(A)$
		Inverso di A_λ non continuo	spettro continuo	$\sigma_c(A)$
	Codominio di A_λ non denso in S		spettro residuo	$\sigma_r(A)$
A_λ non invertibile			spettro puntuale	$\sigma_p(A)$

Dallo schema "si leggono" le definizioni seguenti:

Spettro continuo, $\sigma_c(A)$ = insieme dei $\lambda \in C$ tali che:

A_λ invertibile, codominio denso in S, inverso non continuo

Spettro residuo, $\sigma_r(A)$ = insieme dei $\lambda \in C$ tali che:

A_λ invertibile e codominio non denso in S

Spettro puntuale, $\sigma_p(A)$ = insieme dei $\lambda \in C$ tali che:

A_λ non invertibile.

Introduciamo ora alcune locuzioni relative allo spettro puntuale.

Ogni numero λ appartenente allo spettro puntuale si chiama un autovalore di A. Dunque λ è un autovalore di A quando il nucleo di A_λ non è costituito dal solo 0, e ciò significa che l'equazione nell'incognita x:

$$Ax = \lambda x \tag{2}$$

ammette soluzioni diverse da quella banale $x = 0$. Ogni soluzione $x \neq 0$ della (2) si chiama un autovettore dell'operatore A; più precisamente, un autovettore di A corrispondente all'autovalore λ.

Se λ è un autovalore di A, il nucleo di A_λ, ossia il sottospazio di S costituito dallo 0 di S e dagli autovettori corrispondenti all'autovalore λ, si chiama l'autospazio di A corrispondente all'autovalore λ. Se il nucleo di A_λ ha dimensione finita k (con $k \neq 0$ perché λ è autovalore), come accade ad es. quando S è uno spazio normato ed A è un operatore compatto [cfr. (5.2) pag. 166], il numero k si chiama la molteplicità dell'autovalore λ, e si dice che l'autovalore λ è multiplo di ordine k; se il nucleo di A_λ ha dimensione infinita, si dice che l'autovalore λ ha molteplicità infinita.

Quando lo spazio S ha dimensione finita n, l'operatore lineare A si identifica con una matrice quadrata di ordine n, e perciò gli autovalori di A si chiamano autovalori della matrice. Si veda quanto riportato in proposito a pag. 251 bis.

2. Insieme risolvente e spettro di un operatore lineare e continuo di uno spazio di Banach in sè. Raggio spettrale.

Nel caso che S sia uno spazio di Banach e l'operatore lineare $A: S \to S$ sia continuo, le condizioni (1) che esprimono l'appartenenza di λ all'insieme

risolvente di A si esprimono più concisamente. Infatti è facile verificare che, nelle predette ipotesi su S ed A, un numero λ appartiene all'insieme risolvente di A se e solo se è verificata la condizione:

A_λ è dotato di inverso definito in S

cioè:

$$(3)\quad \begin{cases} a)\ A_\lambda \text{ invertibile} \\ b')\ (\text{codominio di } A_\lambda) = S \end{cases} .$$

Dim. Se valgono le (1), l'operatore lineare A è un isomorfismo di S sul suo codominio $A(S)$, e quindi tale codominio è chiuso [Teor. (2.3) pag. 140]; poichè, insieme $A(S)$ chiuso in S, risulta $A(S)=S$ e quindi sussistono le (3).
Viceversa, se valgono le (3), delle (1) la a) e la b) sussistono banalmente, mentre la c) consegue dal teorema di Banach [Teor. (2.4) pag. 141].

Nelle considerazioni che seguono supponiamo che S sia uno spazio di Banach e che l'operatore lineare A sia continuo, conformemente all'oggetto di questo paragrafo.

Si dimostra, ma noi ci limitiamo a riportare l'enunciato, che:

(2.1) L'insieme risolvente di A è aperto e non vuoto. Lo spettro di A, conseguentemente chiuso, è anch'esso non vuoto.

Il fatto che $\rho(A)\neq\emptyset$ ci è già noto: abbiamo visto [cfr. (3.3) pag. 216] che per $|\lambda|>\|A\|$ l'operatore A_λ è invertibile e il suo inverso, definito in tutto S, si può esprimere come somma di una serie:

$$(4)\qquad A_\lambda^{-1} = -\frac{1}{\lambda}\left(I+\frac{A}{\lambda}+\frac{A^2}{\lambda^2}+\cdots+\frac{A^{n-1}}{\lambda^{n-1}}+\cdots\right)$$

Di conseguenza, detto $\Gamma(r)$ il cerchio chiuso del piano complesso di centro 0 e raggio r, i punti λ esterni al cerchio $\Gamma(\|A\|)$ appartengono tutti all'insieme risolvente $\rho(A)$, e lo spettro di A è contenuto in tale cerchio.

La precedente conclusione si può perfezionare. Infatti si può dimostrare che esiste il limite:

$$(5)\qquad \lim_n \sqrt[n]{\|A^n\|} \leq \|A\| ,$$

diciamolo r_A, e che i punti esterni al cerchio $\Gamma(r_A)$, contenuto in $\Gamma(\|A\|)$, appartengono tutti all'insieme risolvente di A; di conseguenza lo spettro di A è contenuto nel cerchio $\Gamma(r_A)$. Inoltre la (4) sussiste non solo per $|\lambda|>\|A\|$, ma per tutti i λ tali che $|\lambda|>r_A$ (il risultato è ovviamente significativo se $r_A<\|A\|$).

Se λ è un punto di $\Gamma(r_A)$, naturalmente non è detto che λ appartenga allo spettro di A, ossia può accadere che appartenga all'insieme risolvente di

A. Notiamo che se S ha dimensione infinita, e se l'operatore A è compatto, il punto 0 appartiene certamente allo spettro [cfr. es.4 pag.170].

Il numero r_A, ossia il limite (5), si chiama il raggio spettrale dell'operatore A.

3. Caso di uno spazio di Hilbert.

Importanti risultati sussistono nel caso che S sia uno spazio di Hilbert, sempre con riferimento ad un operatore $A: S \to S$ lineare e continuo. Ne riportiamo alcuni, riassumendoli in tre enunciati: nel primo si suppone A autoaggiunto, nel secondo A compatto, nel terzo A autoaggiunto e compatto.

(3.1). Se S è uno spazio di Hilbert e se l'operatore *lineare e continuo* $A: S \to S$ è *autoaggiunto* si ha:

1) Lo spettro di A è costituito da numeri reali, ed ha per minimo e massimo rispettivamente gli estremi $m(A)$, $M(A)$ dell'operatore A;

2) lo spettro residuo di A è vuoto;

3) due autovettori corrispondenti ad autovalori distinti sono ortogonali.

Dimostrazione.

1) Si tratta di dimostrare che ogni numero λ *non* reale, oppure reale *ma* esterno all'intervallo $[m(A), M(A)]$, appartiene all'insieme risolvente di A. Per semplicità scriveremo m, M invece di $m(A)$, $M(A)$. Per ogni $x \in S$, essendo

$$(A_\lambda x, x) = (Ax - \lambda x, x) = (Ax, x) - \lambda \|x\|^2$$

si ha :

$$\left|(Ax, x) - \lambda \|x\|^2\right| \leq \|A_\lambda x\| \cdot \|x\| \tag{6}$$

Minoriamo il primo membro della (6).

- Per λ *reale*, risultando $m\|x\|^2 \leq (Ax, x) \leq M\|x\|^2$ (per la definizione stessa di m ed M), si ha:

$$\text{se } \lambda < m, \quad \left|A(x,x) \quad \lambda\|x\|^2\right| = (Ax, x) \cdot \ \lambda\|x\|^2 \geq (m \quad \lambda)\|x\|^2$$

$$\text{se } \lambda > M, \quad \left|A(x,x) - \lambda\|x\|^2\right| = \lambda\|x\|^2 - (Ax, x) \geq (\lambda - M)\|x\|^2 .$$

- Per λ *non reale*, essendo $\mathrm{Im}(Ax, x) = 0$, [cfr.(4.7) pag.222 BIS][1], si ha :

$$\left|A(x,x) - \lambda\|x\|^2\right| \geq \left|\mathrm{Im}\left[(Ax, x) - \lambda\|x\|^2\right]\right| = |\mathrm{Im}(\lambda)| \cdot \|x\|^2 .$$

In ogni caso, se λ *non* è reale, oppure è reale *ma* non appartiene all'intervallo $[m, M]$, esiste una costante positiva c^+ tale da risultare:

$$c^+\|x\|^2 \leq \left|(Ax, x) - \lambda\|x\|^2\right| ,$$

e quindi dalla (6) consegue:

$$c^+\|x\| \leq \|A_\lambda x\| \quad \forall x \in S . \tag{7}$$

Per il teorema (2.2) di pag.139 se ne deduce che A_λ è invertibile. Per dimostrare che λ appartiene all'insieme risolvente di A basta allora dimostrare che $A_\lambda(S) = S$.

A tale scopo osserviamo intanto che la (7) sussiste anche con $\overline{\lambda}$ in luogo di λ, sicché per il teorema dianzi citato anche $A_{\overline{\lambda}}$ ossia A^*_λ è invertibile, ed ha l'inverso continuo. Dunque A^*_λ è un isomorfismo dello spazio S sul sottospazio $A^*_\lambda(S)$: ne segue che $A^*_\lambda(S)$ è chiuso [(2.3) pag.140][2]. Per il teor.(4.12) di pag.226 con $B = A^*$, dal fatto che A_λ è invertibile e che l'aggiunto ha il codominio chiuso, consegue che $A_\lambda(S) = S$.

---

[1] Sia al solito $A: H \to S$ un operatore lineare, con S spazio di Hilbert ed H sottospazio vettoriale di S denso in S. Se lo spazio S è complesso, sussiste l'equivalenza: A autoaggiunto $\Leftrightarrow (Ax, x)$ reale $\forall x \in H$.

[2] Se l'operatore lineare A di S in S_1 è un isomorfismo di S su $A(S)$, e se S è uno spazio di Banach, il sottospazio $A(S)$ di S_1 è chiuso in S_1.

Resta da far vedere che i numeri m ed M appartengono allo spettro di A. Essendo A autoaggiunto, la funzione:

$$b(x,y) = (Ax - mx, y)$$

è una forma sesquilineare simmetrica[3] in $S \times S$, e si ha

$$b(x,x) = (Ax - mx, x) = (Ax, x) - m\|x\|^2 \geq 0\,.$$

Di conseguenza risulta per la disuguaglianza di Schwartz[4]:

$$'b(x,y)| \leq \sqrt{b(x,x)}\sqrt{b(y,y)}$$

ossia

$$|(Ax - mx, y)| \leq \sqrt{(Ax - mx, x)}\sqrt{(Ay - my, y)}\,,$$

e quindi

$$|(Ax - mx, y)| \leq \sqrt{(Ax - mx, x)}\sqrt{\|Ay - my\| \cdot \|y\|}\,. \tag{8}$$

D'altra parte si ha:

$$\|Ay - my|^{\cdot} = \ddot{}(A - mI)y\| \leq \|(A - mI)\| \cdot \|y\|\,,$$

e quindi, ponendo nella (8) $y = Ax - mx$, si ottiene:

$$\ddot{}Ax - mx\|^2 \leq \sqrt{(Ax - mx, x)}\sqrt{\ddot{}A - mI\| \cdot \|Ax - mx\|^2}$$

da cui segue

$$\|Ax - mx\| \leq \sqrt{\|A - mI\|}\sqrt{(Ax - mx, x)}$$

e quindi

$$\|Ax - mx\ddot{}\ddot{} \leq \sqrt{\ddot{}A - mI\|}\sqrt{(Ax, x) - m\|x\|^2}\,. \tag{9}$$

Denotiamo ora con $\{x_n\}$ una successione di punti di S tali che[5]

$$\|x_n\| = 1, \quad \lim(Ax_n, x_n) = m\,. \tag{10}$$

In virtù della (9) si ha:

$$\lim(Ax_n - mx_n) = 0 \tag{11}$$

e da ciò segue che m appartiene allo spettro di A. Infatti se m appartenesse all'insieme risolvente di A, dalla banale relazione

$$x_n = (A - mI)^{-1}(Ax_n - mx_n)$$

per la (11) si avrebbe $\lim x_n = 0$, in contrasto con la prima delle (10).

Analogamente si dimostra che M appartiene allo spettro di A, considerando questa volta però la forma bilineare $b(x,y) = (Mx - Ax, y)$.

2) Ammettiamo che esista un λ appartenente allo spettro residuo. Siccome $A_\lambda(S)$ *non* è denso in S, per la 4.6 di pag.121[6] risulta

$$A_\lambda(S)^\perp \neq \{0\}$$

---

[3] Infatti si ha: $b(y,x) = (Ay - my, x) = (Ay, x) \cdot \ m(y,x) = (y, Ax) \ \cdot m(y,x) = \overline{(Ax, y)} \quad m\overline{(x,y)} = \overline{b}(x,y)$.

[4] Nella dimostrazione della disuguaglianza di Schwartz [cfr. pag.94], la proprietà del prodotto scalare $p(x,x) > 0\ \forall x \in S - \{0\}$ si utilizza solo in quanto implica $p(x,x) \geq 0\ \forall x \in S$, e quindi la disuguaglianza stessa è valida anche per la forma bilineare $b(x,y)$.

[5] Essendo $m(A) = \inf_{\|x\|=1}(Ax, x)$, il punto m è aderente all'insieme costituito dai numeri reali (Ax, x) con $\|x\| = 1$, pertanto esiste una successione $\{x_n\}$ soddisfacente la (10).

[6] Per ogni sottospazio vettoriale H dello spazio di Hilbert S, sussiste l'equivalenza

$$\overline{H} = S \ \Leftrightarrow \ H^\perp = \{0\}\,;$$

in altri termini, se non esiste alcun vettore diverso dal vettore nullo che sia ortogonale ad H, allora H è denso in S.

e quindi esiste un $y \in S - \{0\}$ ortogonale ad $A_\lambda(S)$: allora y appartiene al nucleo dell'aggiunto di A [cfr.(17) pag.226], che è lo stesso A_λ in quanto A è autoaggiunto e λ è reale per la 1) già dimostrata. Pertanto λ appartiene allo spettro puntuale di A (esistendo elementi non nulli di $\ker A_\lambda$), e ciò è assurdo perché λ appartiene allo spettro residuo.

3) Sia $Ax = \lambda_1 x$, $Ay = \lambda_2 y$, con $x \neq 0$, $y \neq 0$, $\lambda_1 \neq \lambda_2$. Si ha

$$\lambda_1(x,y) = (\lambda_1 x, y) = (Ax, y) = (x, Ay) = (x, \lambda_2 y) = \lambda_2(x,y)$$

da cui segue

$$[\lambda_1 - \lambda_2](x,y) = 0$$

e quindi $(x,y) = 0$.

— — — — — — — — — —

(3.2). Se S è uno spazio di Hilbert e se l'operatore lineare e continuo $A: S \to S$ è compatto, si ha:

1) Ogni autovalore $\lambda \neq 0$ ha molteplicità finita,

2) L'insieme degli autovalori di A (spettro puntuale) è finito o numerabile, e nel secondo caso ha come unico punto di accumulazione lo zero.

3) Lo spettro continuo di A o è vuoto, o è costituito dall'unico punto 0.

— — — — — — — — — —

Dim. La 1) consegue immediatamente dalla (5.2) di pag.166.

Della 2) rinunciamo ad esporre la dimostrazione.

Dimostriamo la 3). Per il teorema dell'alternativa [(6.1) pag.233], un qualunque numero complesso $\lambda \neq 0$ o appartiene all'insieme risolvente di A oppure è autovalore di A. Ossia, i numeri $\lambda \neq 0$ non possono appartenere allo spettro continuo: pertanto allo spettro continuo può appartenere solo il punto 0.

— — — — — — — — — —

(3.3). Se S è uno spazio di Hilbert e se l'operatore lineare e continuo $A: S \to S$ è compatto e autoaggiunto, oltre alle conclusioni di (3.1), (3.2) si ha che:

— Se $A \neq 0$, almeno uno dei numeri $\|A\|, -\|A\|$ è autovalore di A.

— — — — — —

Dim. In virtù della (16) di pag.233, o risulta $\|A\| = m(A)$ (e quindi vale almeno una delle uguaglianze $\|A\| = m(A)$, $\|A\| = -m(A)$), o risulta $\|A\| = |M(A)|$ (e quindi vale almeno una delle uguaglianze $\|A\| = M(A)$, $\|A\| = -M(A)$). Nel primo caso, uno dei numeri $\|A\|, -\|A\|$ è uguale ad $m(A)$, nel secondo caso uno di tali numeri è uguale ad $M(A)$. In ogni caso, per la 1) di (3.1), almeno uno dei numeri $\|A\|, -\|A\|$ appartiene allo spettro di A; per la 2) di (3.1) non può appartenere allo spettro residuo, e per la 3) di (3.2) non può appartenere allo spettro continuo. Dunque almeno uno dei predetti numeri è autovalore di A.

— — — — — — — — — — — —

Supposto che S sia uno spazio di Hilbert, riordiniamo le conclusioni sulla decomposizione dello spettro di un operatore compatto e autoaggiunto di S in S, nonché le eventualità che possono presentarsi in corrispondenza di un dato numero complesso λ.

(3.4). Siano: S uno spazio di Hilbert, A un operatore lineare di S in S, compatto e autoaggiunto.

Lo spettro residuo di A è vuoto.

Lo spettro continuo di A o è vuoto, oppure è $\{0\}$.

Lo spettro puntuale di A (insieme degli autovalori) è non vuoto, finito o numerabile, costituito da numeri reali appartenenti all'intervallo $[m(A), M(A)]$ i cui estremi ne sono il minimo e il massimo; se $A \neq 0$, almeno uno di tali estremi coincide o con $\|A\|$ o con $-\|A\|$.

Quando è numerabile, lo spettro puntuale di A ammette un unico punto di accumulazione, che è 0.

Gli autovalori non nulli di A sono di molteplicità finita, e gli autospazi corrispondenti sono a due a due ortogonali.

(3.5). Siano: S uno spazio di Hilbert, A un operatore lineare di S in S, compatto e autoaggiunto. Per ogni numero complesso λ si ha quanto segue:

Se λ non è reale, λ appartiene all'insieme risolvente di A.

Se λ è reale diverso da 0, λ o appartiene all'insieme risolvente o è autovalore di A.

Il punto $\lambda = 0$ o appartiene allo spettro continuo o è autovalore di A; se lo spazio S ha dimensione finita, lo zero può appartenere anche all'insieme risolvente.

4. Rappresentazione spettrale di un operatore compatto e autoaggiunto in uno spazio di Hilbert. Operatori del tipo f(A), con f funzione reale ed A operatore.

Se A è un operatore lineare di uno spazio di Hilbert in sè, compatto e autoaggiunto, gli autovalori di A costituiscono o un insieme finito, oppure un insieme numerabile avente come unico punto di accumulazione lo zero: le considerazioni che seguono si riferiscono al secondo caso, quindi ad uno spazio S di dimensione infinita (*).

Rappresentiamo gli autovalori di A diversi da 0 mediante una successione $\{\lambda_n\}$ tale che la successione dei moduli sia decrescente, ed inoltre

(*) Tale conseguenza sarà ovvia tra breve.

tale che ciascun autovalore di molteplicità $k>1$ sia ripetuto k volte (ad es. se λ_1 ha molteplicità $k>1$ nella successione $\{\lambda_n\}$ sarà $\lambda_1=\lambda_2=\cdots=\lambda_k$).

Ovviamente risulta:

$$\lim \lambda_n = 0 , \tag{11}$$

e alla successione:

$$\lambda_1 , \lambda_2 , \dots , \lambda_n , \dots \tag{12}$$

si può associare una successione ortonormale di elementi di S:

$$x_1 , x_2 , \dots , x_n , \dots \tag{13}$$

in modo che per ogni $n\in N$ l'elemento x_n sia un autovettore di A corrispondente all'autovalore λ_n; a tale scopo, ricordando che autovettori corrispondenti ad autovalori distinti sono ortogonali, basta scegliere per ciascun autovalore di molteplicità 1 un autovettore di norma unitaria, e per ciascun autovalore di molteplicità $k>1$ altrettanti autovettori di norma unitaria a due a due ortogonali (*). Naturalmente gli elementi delle successioni (12) e (13) sono legati dalla relazione:

$$Ax_n = \lambda_n x_n \quad \forall n\in N . \tag{14}$$

Per ogni $n\in N$ denotiamo poi con P_n il proiettore di S sul sottospazio generato dal vettore x_n [cfr. (11) pag. 118]:

$$P_n : x\in S \to (x,x_n)x_n . \tag{15}$$

Del teorema che segue non riportiamo la dimostrazione; proveremo invece che la tesi del teorema si può esprimere in diverse altre forme.

(4.1). Teorema di Hilbert-Schmidt.

Siano S uno spazio di Hilbert ed A un operatore lineare di S in S, compatto e autoaggiunto.

Se $\{\lambda_n\}$ è la successione (12), e P_n è il proiettore (15) con x_n termine n^{mo} della (13), risulta:

$$A = \sum_{n\geq 1}^{\infty} \lambda_n P_n \tag{16}$$

la convergenza essendo puntuale in senso forte; si ha cioè:

$$Ax = \sum_{n\leq}^{\infty} \lambda_n (x,x_n)x_n \quad \forall x\in S . \tag{17}$$

(*) In altri termini, basta scegliere per ciascun autovalore una base ortonormale dell'autospazio corrispondente.

La (16) prende il nome di <u>decomposizione spettrale dell'operatore A</u>.

Dimostriamo che:

(4.2) <u>Nelle ipotesi del teorema (4.1), la tesi è equivalente a ciascuna delle proposizioni seguenti:</u>

1) <u>Per ogni $x \in S$ esiste un elemento x_0 del nucleo di A tale che:</u>

(18) $$x = x_0 + \sum_{n=1}^{\infty} (x, x_n) x_n .$$

<u>Di conseguenza il sistema ortonormale $\{x_n\}$ è completo se e solo se l'operatore A è invertibile, cioè se e solo se 0 non è autovalore di A.</u>

2) <u>Qualunque siano $x, y \in S$, risulta:</u>

(19) $$(Ax, y) = \sum_{n=1}^{\infty} \lambda_n (x, x_n)(x_n, y) .$$

3) <u>Per ogni $x \in S$, l'elemento Ax è la somma della sua serie di Fourier rispetto al sistema ortonormale $\{x_n\}$, cioè risulta:</u>

(20) $$Ax = \sum_{n=1}^{\infty} (Ax, x_n) x_n .$$

Se 0 non è autovalore di A, la (18) fornisce:

(21) $$x = \sum_{n=1}^{\infty} (x, x_n) x_n \qquad \forall x \in S,$$

cioè, considerando il proiettore (15) e denotando con I l'identità in S:

(21') $$I = \sum_{n=1}^{\infty} P_n .$$

(<u>rappresentazione spettrale dell'identità, associata all'operatore A</u>)

Dim. (17) $\Rightarrow$ 1). La serie a secondo membro della (18) è la serie di Fourier di x rispetto al sistema ortonormale $\{x_n\}$: detta w la sua somma, il vettore x-w è ortogonale a tutti i vettori x_n [(6.4) pag.128 e (4.3) pag.118], e quindi scrivendo la (17) con x-w in luogo di x si ottiene:

$$A(x-w) = \sum_{n=1}^{\infty} \lambda_n (x-w, x_n) x_n = 0 .$$

Pertanto $x-w \in \ker A$ e quindi, posto $x-w = x_0$, sussiste la (18).

Resta da provare la conseguenza formulata dopo la (18).

Se A è invertibile, ossia $\ker A = \{0\}$, il vettore x_0 che figura nella (18) è nullo qualunque sia $x \in S$. Dunque ogni $x \in S$ è la somma della sua serie di Fourier rispetto ad $\{x_n\}$, e ciò prova che il sistema ortonormale $\{x_n\}$ è completo [(6.5) pag.129].

Viceversa, supposto $\{x_n\}$ completo, e detto x un arbitrario elemento di $\ker A$, per la (17) si ha:

$$0 = \sum_{n=1}^{\infty} \lambda_n (x, x_n) x_n$$

da cui, moltiplicando scalarmente per x_i, si deduce che x è ortogonale a ciascuno dei vettori x_i: per la completezza di $\{x_n\}$ ciò comporta che $x = 0$ [pag.124]. Dunque $\ker A = \{0\}$, ossia A è invertibile.

Dim. 1) $\Rightarrow$ 2). Basta far vedere che 1) $\Rightarrow$ (17), perché dalla (17) moltiplicando scalarmente per y si ottiene la (19).

Orbene, dalla (18) consegue, tenendo presente la (14):

$$Ax = Ax_0 + A\left[\sum_{i=1}^{\infty}(x,x_i)x_i\right] = 0 + A\left(\lim_{n\to\infty}\sum_{i=1}^{n}(x,x_i)x_i\right) = \lim_{n\to\infty} A\left(\sum_{i=1}^{n}(x,x_i)x_i\right) =$$

$$= \lim_{n\to\infty}\sum_{i=1}^{n}(x,x_i)Ax_i = \lim_{n\to\infty}\sum_{i=1}^{n}(x,x_i)\lambda_i x_i = \sum_{i=1}^{\infty}\lambda_i(x,x_i)x_i$$

ossia la (17).

Dim. 2) $\Rightarrow$ 3). La serie a secondo membro della (20) è convergente, perché è la serie di Fourier di un elemento di S rispetto al sistema ortonormale $\{x_n\}$ [(6.4) pag. 178]. Detta z la sua somma, cioè posto:

$$z = \sum_{n=1}^{\infty}(Ax,x_n)x_n,$$

per ogni $y\in S$ si ha, tenuto conto che A è autoaggiunto e che sussiste la (14):

$$(22)\qquad (z,y) = \sum_{n=1}^{\infty}(Ax,x_n)(x_n,y) = \sum_{n=1}^{\infty}(x,Ax_n)(x_n,y) = \sum_{n=1}^{\infty}\lambda_n(x,x_n)(x_n,y),$$

e quindi, per la (19), risulta $(z,y) = (Ax,y)$. Essendo $(z-Ax,y)=0$ $\forall y\in S$, si ha $z=Ax$ e ciò prova la (20).

Dim. 3) $\Rightarrow$ (17). Le (22), senza il primo membro, provano che il secondo membro della (20) è uguale al secondo membro della (17).

Una conseguenza del teorema di Hilbert-Schmidt è il seguente teorema:

(4.3) Siano: S uno spazio di Hilbert, A un operatore lineare di S in S, compatto e autoaggiunto, λ un numero complesso non nullo e non autovalore di A.

Qualunque sia $y\in S$, la soluzione dell'equazione:

$$(23)\qquad Ax - \lambda x = y \qquad (1)$$

è data dalla formula:

$$(24)\qquad x = -\frac{1}{\lambda}y + \frac{1}{\lambda}\sum_{n=1}^{\infty}\frac{\lambda_n}{\lambda_n-\lambda}(y,x_n)x_n$$

dove $\{\lambda_n\}, \{x_n\}$ sono le successioni (12), (13).

Se 0 non è autovalore di A, la (24) può scriversi:

$$(25)\qquad x = \sum_{n=1}^{\infty}(\lambda_n-\lambda)^{-1}(y,x_n)x_n .$$

---

(1) Per la (3.5) di pag. 265 il numero λ appartiene all'insieme risolvente di A.

Dim. Moltiplicando scalarmente per x_n ambo i membri della (23) si ha:

$$(Ax, x_n) - \lambda (x, x_n) = (y, x_n)$$

ossia, essendo A autoaggiunto:

$$(x, Ax_n) - \lambda (x, x_n) = (y, x_n)$$

e quindi, per la (14) di pag. 246:

$$(\lambda_n - \lambda)(x, x_n) = (y, x_n),$$

da cui segue:

$$(x, x_n) = \frac{1}{\lambda_n - \lambda} (y, x_n).$$

Sostituendo nella (17) [teor. di Hilbert-Schmidt] si ottiene:

$$Ax = \sum_{n=1}^{\infty} \frac{\lambda_n}{\lambda_n - \lambda} (y, x_n) x_n$$

da cui, sostituendo nella (23), si deduce la (24).

Se 0 non è autovalore di A, per la (21) risulta:

$$y = \sum_{n=1}^{\infty} (y, x_n) x_n$$

e la (24) diviene:

$$x = \frac{1}{\lambda} \sum_{n=1}^{\infty} \left[-1 + \frac{\lambda_n}{\lambda_n - \lambda} \right] (y, x_n) x_n ,$$

e questa uguaglianza è la (25).

Nelle ipotesi della (4.3), supponiamo che 0 non sia autovalore di A, e conveniamo di denotare con $A - \lambda$ l'operatore $A - \lambda I$.

Detto P_n il proiettore di S sul sottospazio generato dall'autovettore x_n [cfr. (15) pag. 246], osserviamo le seguenti uguaglianze:

$$A = \sum_{n=1}^{\infty} \lambda_n P_n \quad (1), \qquad A^k = \sum_{n=1}^{\infty} \lambda_n^k P_n \quad (k \in N) \quad (2)$$

$$A - \lambda = \sum_{n=1}^{\infty} (\lambda_n - \lambda) P_n \quad (3), \qquad (A - \lambda)^{-1} = \sum_{n=1}^{\infty} (\lambda_n - \lambda)^{-1} P_n \quad (4).$$

Le serie a secondo membro sono tutte del tipo:

$$\sum_{n=1}^{\infty} f(\lambda_n) P_n , \tag{26}$$

e l'operatore a primo membro si ottiene formalmente dall'espressione $f(\lambda_n)$ sostituendo in essa λ_n con A. Ciò suggerisce di denotare con $f(A)$ l'operatore (26) qualunque

(1) E' la (16) di pag. 246

(2) Consegue facilmente dalla (17) di pag. 246, sostituendovi x con Ax e procedendo poi induttivamente.

(3) Si deduce dalla prima uguaglianza e dalla (21') di pag. 247.

(4) E' la (25) di pag. 248, in quanto il primo membro di detta (25) è $(A - \lambda)^{-1}$.

sia la funzione reale f di una variabile reale, definita nell'intervallo $[m(A), M(A)]$ [1] ed ivi continua.

Ciò posto, per ogni funzione f siffatta si pone (con A operatore lineare invertibile compatto autoaggiunto):

$$f(A) = \sum_{n=1}^{\infty} f(\lambda_n) P_n \tag{27}$$

la convergenza intende puntuale in senso forte. La serie a secondo membro è convergente puntualmente in senso forte, in quanto per ogni $x \in S$ la serie di termine n^{mo}:

$$f(\lambda_n) P_n x = f(\lambda_n)(x, u_n) x_n$$

è convergente in virtù della (6.2) di pag. 125 [2].

Si dimostra che anche l'operatore $f(A)$ è autoaggiunto, e che qualunque sia la funzione reale g continua nell'intervallo $[m(A), M(A)]$ risulta:

$$f \cdot g\,(A) = f(A)\, g(A) . \tag{28}$$

In particolare, se $f \neq 0$ l'operatore $f(A) \cdot \frac{1}{f}(A)$ è l'operatore identico di S in S, sicché $f(A)$ è invertibile e l'operatore inverso è $\frac{1}{f}(A)$.

Ad esempio si possono considerare gli operatori $\log A$, e^A, $\sqrt{A}$, $\cos A$, ecc.: il primo nell'ipotesi che A sia strettamente positivo, il terzo nell'ipotesi che A sia positivo. Notiamo che dalla (28) con $f(t) = g(t) = \sqrt{t}$ consegue:

$$\sqrt{A} \cdot \sqrt{A} = A$$

Aggiungiamo infine, senza dimostrazione, che se $f(t) \geq 0 \ \forall t \in [m(A), M(A)]$ risulta $f(A) \geq 0$, e che la norma dell'operatore $f(A)$ verifica la relazione:

$$\| f(A) \| \leq \max_{t \in [m,M]} |f(t)|$$

avendo indicato con m, M gli estremi dell'operatore A.

I risultati di questo numero si estendono agli operatori autoaggiunti non necessariamente compatti. Di tale estensione daremo un cenno nell'Appendice [pag. 306].

---

(1) Ciò perché gli autovalori λ_n appartengono a tale intervallo, e deve avere senso $f(\lambda_n)$.

(2) I termini della serie costituiscono una successione ortogonale. La successione dei coefficienti appartiene ad l^2, in quanto la serie

$$\sum_{n=1}^{\infty} |f(\lambda_n)(x, u_n)|^2$$

è maggiorata dalla serie:

$$(\max |f|)^2 \sum_{n=1}^{\infty} |(x, u_n)|^2$$

che è convergente per la disuguaglianza di Bessel [cfr. (10) pag. 121].

Una importante applicazione della teoria spettrale e' costituita dagli operatori integrali, e dalle equazione integrale di Fredholm introdotta nell'esempio 1 di pag. 168, ripresa poi nell'esempio di pag. 236.

A tale applicazione erano dedicate la presente pagina 251, e le pagine 252-253. L'esposizione e' stata successivamente ampliata, e perciò trasferita nell'Appendice [si veda pag. 335].

Pertanto le predette pagine 252-253 sono state soppresse, e nella presente pagina 251 figura una breve aggiunta all'esercizio sulle equazioni integrali svolto a pag. 237.

* * *

Nel caso dell'esercizio di pag. 237 gli autovalori dell'operatore J_1 sono due, e sono dati dalla (21) di pag. 238.

Gli autovettori corrispondenti (che si chiamano anche autosoluzioni) si ottengono risolvendo il sistema (19) di pag. 237 prima con $\lambda=\lambda_1$ e poi con $\lambda=\lambda_2$, essendo λ_1, λ_2 dati dalla (21); si pone poi $x(s)=as+b$, essendo (a,b) la soluzione del sistema, prima con $\lambda=\lambda_1$ e poi con $\lambda=\lambda_2$.

Si trova:

$$x_1(s)=(1+s\sqrt{3})\,h \quad , \qquad x_2(s)=(1-s\sqrt{3})\,h \ ,$$

essendo h un'arbitraria costante non nulla.

La rappresentazione spettrale dell'operatore J_1 e' data dall'uguaglianza:

$$J_1x(s)=\lambda_1(x,x_1)x_1(s)+\lambda_2(x,x_2)x_2(s) \qquad \forall x\in L^2(\Omega)$$

cioe' dall'uguaglianza:

$$J_1x(s)=\frac{1+s\sqrt{3}}{2\sqrt{3}}\int_0^1 x(t)(1+t\sqrt{3})\,dt-\frac{1-s\sqrt{3}}{2\sqrt{3}}\int_0^1 x(t)(1-t\sqrt{3})\,dt \qquad \forall x\in L^2(\Omega).$$

Il lettore osservi che, trasportando sotto il segno di integrale ciascuno dei due fattori, e scrivendo l'integrale della differenza in luogo della differenza degli integrali, si ritrova l'espressione di J_1x che figura a primo membro della (17) di pag. 237, ivi implicitamente indicata con Jx.

Aggiunta al n. 1 di pag. 239.

Quando lo spazio S ha dimensione finita n, l'operatore A si identifica con una matrice quadrata di ordine n. Gli autovalori di A, che si chiamano anche autovalori della matrice, sono allora soluzioni di un'equazione algebrica di grado n.

Vogliamo qui segnalare che la molteplicità di un autovalore λ_0 non coincide necessariamente con la molteplicità di λ_0 come radice della predetta equazione algebrica, che si chiama <u>la molteplicità algebrica</u> dell'autovalore λ_0.

Riportiamo in proposito il seguente esempio.

Consideriamo l'operatore lineare:

$$A: (x_1, x_2) \in \mathbb{R}^2 \to (2x_1 + 2x_2 , -\tfrac{1}{2} x_1) \in \mathbb{R}^2$$

ossia l'operatore $A: \mathbb{R}^2 \to \mathbb{R}^2$ definito ponendo:

$$A(x_1, x_2) = \begin{bmatrix} 2 & 2 \\ -\frac{1}{2} & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \qquad \forall (x_1, x_2) \in \mathbb{R}^2 .$$

Posto $x = (x_1, x_2)$, per ogni $\lambda \in \mathbb{C}$ si ha:

$$Ax - \lambda x = \left((2-\lambda) x_1 + 2x_2 , -\tfrac{1}{2} x_1 - \lambda x_2 \right)$$

e quindi $\ker (A - \lambda I)$ è l'insieme delle soluzioni del sistema:

$$\begin{cases} (2-\lambda) x_1 + 2 x_2 = 0 \\ -\frac{1}{2} x_1 - \lambda x_2 = 0 \end{cases} .$$

Gli autovalori di A sono le soluzioni dell'equazione algebrica:

$$\begin{vmatrix} 2-\lambda & 2 \\ -\frac{1}{2} & -\lambda \end{vmatrix} = 0 \quad \text{ossia} \quad (\lambda - 1)^2 = 0 .$$

Tale equazione ammette dunque l'unica radice $\lambda = 1$ di molteplicità 2.

L'autospazio relativo a $\lambda = 1$ è l'insieme delle soluzioni del sistema:

$$\begin{cases} x_1 + 2x_2 = 0 \\ \frac{1}{2} x_1 + x_2 = 0 \end{cases}$$

che sono le coppie $(-2t, t)$ con $t \in \mathbb{R}$. Il sottospazio di tali coppie è costituito dai vettori di $\mathbb{R}^2$ proporzionali al vettore $(-2, 1)$: dunque l'autospazio relativo all'autovalore $\lambda = 1$ ha dimensione 1, sicché l'autovalore ha molteplicità 1. Invece la molteplicità algebrica è 2.

Capitolo IX

Disequazioni variazionali ed applicazioni.

1. Disequazioni variazionali.

Siano: S uno spazio di Hilbert reale, K un sottoinsieme di S non vuoto, chiuso e convesso, w un punto di S.

Il problema di determinare $u \in K$ tale da aversi:

$$(1) \qquad (u, v-u) \geq (w, v-u) \quad \forall v \in K$$

si chiama, come abbiamo detto a pag.116, una disequazione variazionale.

Abbiamo visto che tale problema ammette una ed una sola soluzione, e che equivale al problema di minimizzare nel convesso K il funzionale:

$$(2) \qquad \Phi(v) = \frac{1}{2}\|v\|^2 - (w, v) .$$

L'unica soluzione della (1), o equivalentemente il punto di K che minimizza il funzionale (2), e' la proiezione di w su K, ossia il punto $u \in K$ tale che si abbia:

$$\|w-u\| = \min_{v \in K} \|w-v\| .$$

Ovviamente assegnare il punto w equivale ad assegnare un funzionale lineare e continuo f definito in S, ossia un elemento di S^*, giacche' per il teorema di Riesz [(7.1) pag.177] ogni elemento f di S^* determina univocamente un punto w di S tale che $\langle f, v \rangle = (w, v)$ $\forall v \in S$.

Orbene, il problema espresso dalla (1) e' un caso particolare del seguente. Assegnati una forma bilineare $a(u,v)$ in $S \times S$, un funzionale f lineare e continuo in S, e un sottoinsieme K di S non vuoto, chiuso e convesso, ricercare un elemento $u \in K$ tale che si abbia:

$$(3) \qquad a(u, v-u) \geq \langle f, v-u \rangle \quad \forall v \in K .$$

Tale problema si esprime scrivendo:

(4) $\qquad u \in K : \quad a(u, v-u) \geq \langle f, v-u \rangle \quad \forall v \in K,$

e si chiama una disequazione variazionale.

Sussiste il seguente teorema di Stampacchia:

(1.1). Se la forma bilineare $a(u,v)$ è continua e coercitiva, qualunque sia $f \in S'$ la disequazione variazionale (4) ammette una e una sola soluzione.

Ricordiamo che la forma bilineare $a(u,v)$ si dice coercitiva [pag. 213] quando esiste un numero positivo α tale da risultare:

(5) $\qquad a(v,v) \geq \alpha \|v\|^2 \quad \forall v \in S.$

Dim. Dimostreremo l'esistenza nell'ulteriore ipotesi che la forma bilineare $a(u,v)$ sia simmetrica: ciò semplifica notevolmente la dimostrazione, in quanto il problema si riconduce allora, come vedremo, al caso che $a(u,v)$ sia il prodotto scalare.

Per stabilire l'unicità, siano u_1, u_2 soluzioni della (4). Per ogni $v \in K$ si ha dunque:

$$a(u_1, v-u_1) \geq \langle f, v-u_1 \rangle$$
$$a(u_2, v-u_2) \geq \langle f, v-u_2 \rangle .$$

Ponendo nella prima $v = u_2$ e nella seconda $v = u_1$ si ottengono le relazioni:

$$a(u_1, u_2-u_1) \geq \langle f, u_2-u_1 \rangle$$
$$-a(u_2, u_2-u_1) \geq -\langle f, u_2-u_1 \rangle$$

che sommate membro a membro forniscono la seguente:

$$a(u_1-u_2, u_2-u_1) \geq 0,$$

ossia $a(u_2-u_1, u_2-u_1) \leq 0$. Tenendo conto di questa, la (5) per $v = u_2-u_1$ implica:

$$\|u_2-u_1\|^2 \leq \frac{1}{\alpha} a(u_2-u_1, u_2-u_1) \leq 0$$

e quindi $u_1 = u_2$.

Se la forma bilineare è simmetrica, risultando per la (5) $a(v,v) > 0$ per ogni $v \in S - \{0\}$ si può assumere $a(u,v)$ come un nuovo prodotto scalare per il sostegno di S. Detto S_1 lo spazio pre-hilbertiano che così si ottiene, e posto:

(6) $$(u,v)_1 = a(u,v)\,,$$

per la (5) risulta:

(7) $$\|v\|^2_{S_1} = (v,v)_1 = a(v,v) \geq \alpha \|v\|^2_S \quad ;$$

d'altra parte, per la continuità di $a(u,v)$ [cfr. (2.1) pag. 209] esiste $\ell > 0$ tale che:

$$|a(u,v)| \leq \ell \|u\|_S \|v\|_S \quad \forall u,v \in S$$

da cui per $u=v$ consegue:

$$a(v,v) \leq \ell \|v\|^2_S \qquad \forall v \in S$$

ossia:

(8) $$\|v\|^2_{S_1} \leq \ell \|v\|^2_S \qquad \forall v \in S\,.$$

Per le (7), (8) si ha dunque:

$$\alpha \|v\|^2_S \leq \|v\|^2_{S_1} \leq \ell \|v\|^2_S \qquad \forall v \in S$$

e pertanto le due norme per il sostegno di S sono equivalenti. Ne segue che anche S_1 è uno spazio di Hilbert, che f è anche elemento di S_1^*, e che l'insieme convesso K è chiuso anche in S_1. Di conseguenza se w è l'elemento di S_1 tale che $\langle f, v\rangle = (w,v)_1$ $\forall v \in S_1$, esiste la proiezione di w su K, diciamola u. L'elemento $u \in K$ soddisfa quindi alla relazione:

(9) $$(w-u, v-u)_1 \leq 0 \quad \forall v \in K \quad (1)$$

ed è l'unico che la soddisfa. La (9) si scrive:

$$(u, v-u)_1 \geq (w, v-u)_1 \quad \forall v \in K$$

ossia:

$$a(u, v-u) \geq \langle f, v-u\rangle \quad \forall v \in K$$

e ciò prova l'asserto.

- - - - - - - - - - - - - -

Si verifica subito che:

(1.2). <u>Se la forma bilineare $a(u,v)$, continua e coercitiva, è anche simmetrica, dire che u è soluzione della disequazione variazionale (4) equivale a dire che u minimizza nel convesso K il funzionale</u>:

(10) $$\Phi(v) = \tfrac{1}{2} a(v,v) - \langle f, v\rangle\,.$$

- - - - - - - - -

Dim. Infatti dire che u è soluzione della (4) equivale a dire che u è soluzione della (9), ossia che u minimizza il funzionale (2) con $\|\cdot\|_1$ e $(\cdot,\cdot)_1$ in luogo di $\|\cdot\|$ e $(\cdot,\cdot)$, nel convesso K. In virtù della (6), ne consegue l'asserto.

(1) Cfr. pag. 116

Nelle ipotesi della (1.2), il fatto che u sia soluzione della disequazione variazionale (4) si può esprimere in un' altra forma considerando, per $v \in S$, il differenziale del funzionale (10) nel punto v, cioè il funzionale lineare e continuo:

$$(11) \qquad d\Phi(v) : z \in S \longrightarrow a(v,z) - \langle f, z \rangle .$$

Vogliamo evidenziare tale circostanza anche se non abbiamo ancora introdotto la nozione di differenziale, che si riferisce a funzionali non necessariamente lineari: perciò tale nozione è anticipata brevemente in una "appendice" inserita nella pagina seguente, 257.1.

Orbene, è evidente che:

(1.3) Se la forma bilineare $a(u,v)$, continua e coercitiva, è anche simmetrica, e se u è un elemento di K, dire che u è soluzione della disequazione variazionale (4) equivale a dire che il differenziale del funzionale (10) nel punto u soddisfa alla condizione:

$$(12) \qquad d\Phi(u)(v-u) \geq 0 \qquad \forall v \in K .$$

Ovviamente nel caso che K sia un sottospazio di S, e perciò chiamato H, la disequazione variazionale (4) equivale al problema di ricercare $u \in H$ tale che si abbia:

$$(13) \qquad a(u,z) = \langle f, z \rangle \qquad \forall z \in H \quad [1].$$

La (13) prende il nome di equazione variazionale. Dal teorema (1.1) discende come caso particolare il seguente teorema di Lax-Milgram:

(1.4). Se la forma bilineare $a(u,v)$ è continua e coercitiva, qualunque sia $f \in S^*$ l'equazione variazionale (13) ammette una e una sola soluzione.

Notiamo che per $a(u,v) = (u,v)$ ed $H = S$, la (1.4) si riduce al teorema di Riesz [cfr. (7.1) pag. 177].

---

(1) Infatti se $K = H$ sottospazio di S, posto ogni $z \in H$ si può porre nella (3) $v = u \pm z$, e si ottiene:

$$a(u,z) \geq \langle f, z \rangle \quad , \quad a(u,-z) \geq \langle f, -z \rangle .$$

Siccome la seconda relazione equivale ad $a(u,z) \leq \langle f, z \rangle$, si ha la (13).
Viceversa se l'elemento u di H verifica la (13), ponendo $z = v - u$ si deduce che è verificata la (3).

APPENDICE I AL n. 1

DIFFERENZIALE SECONDO GATEAUX

Se $f(P)$ è una funzione reale di n variabili reali definita in un aperto Ω ed ivi dotata di tutte le derivate parziali prime, il differenziale di f relativo ad un punto $P_0 \in \Omega$ è il funzionale lineare (necessariamente continuo):

$$(*) \qquad df(P_0): (h_1,\dots,h_n) \in R^n \to \operatorname{grad} f(P_0)\cdot h = (\operatorname{grad} f(P_0), h),$$

dove si è posto $h=(h_1,\dots,h_n)$. Abitualmente si scrive $h = dP = (dx_1,\dots,dx_n)$.

La nozione di differenziale si estende ai funzionali definiti in uno spazio normato S.

Osserviamo intanto che se la predetta funzione f è differenziabile nel punto P_0, detta τ l'asse di versore $\frac{h}{|h|}$ $(h\neq 0)$, si ha:

$$\operatorname{grad} f(P_0)\cdot h = |h|\left(\operatorname{grad} f(P_0)\cdot\frac{h}{|h|}\right) = |h|\frac{\partial f}{\partial \tau} = |h| \lim_{\substack{P\to P_0 \\ P\in\tau}} \frac{f(P)-f(P_0)}{\overline{PP_0}} =$$

$$= |h| \lim_{\tau\to 0} \frac{f\left(P_0+\tau\frac{h}{|h|}\right)-f(P_0)}{\tau} = \lim_{\tau\to 0} \frac{f\left(P_0+\tau\frac{h}{|h|}\right)-f(P_0)}{\tau/|h|} = \lim_{t\to 0}\frac{f(P_0+th)-f(P_0)}{t}$$

e quindi il differenziale $(*)$ si può scrivere:

$$df(P_0): h\in R^n \to \lim_{t\to 0}\frac{f(P_0+th)-f(P_0)}{t}.$$

Orbene, se Φ è un funzionale definito in uno spazio normato S, si dice che Φ è <u>differenziabile</u> (<u>secondo Gateaux</u>, o <u>in senso debole</u>) nel punto $x_0\in S$, quando per ogni $h\in S$ esiste il limite:

$$(**) \qquad \lim_{t\to 0}\frac{\Phi(x_0+th)-\Phi(x_0)}{t},$$

ed inoltre tale limite, come funzione di h, è lineare e continuo.

In tale ipotesi, il funzionale lineare e continuo che ad ogni $h\in S$ associa il limite $(**)$ si denota con $d\Phi(x_0)$, e si chiama <u>il differenziale di Φ nel punto x_0</u>.

È facile verificare che:

1) <u>Ogni funzionale lineare e continuo Φ è differenziabile in ciascun punto di S, e qualunque sia $x_0\in S$ risulta:</u>

$$\langle d\Phi(x_0), h\rangle = \langle \Phi, h\rangle \quad \forall h\in S.$$

2) <u>Se $a(x,y)$ è una forma bilineare in $S\times S$, con S spazio pre-hilbertiano, ed è continua e simmetrica, il funzionale Φ definito ponendo:</u>

$$\Phi(x) = \frac{1}{2}a(x,x) \quad \forall x\in S$$

<u>è differenziabile in ogni punto di S, e qualunque sia $x_0\in S$ risulta:</u>

$$\langle d\Phi(x_0), h\rangle = a(x_0,h) \quad \forall h\in S.$$

Di conseguenza il funzionale (10) di pag. 256 è differenziabile in ogni punto $v\in S$, e il suo differenziale nel punto v è il funzionale (11) di pag. 257.

APPENDICE II AL n.1

Vogliamo qui aggiungere un cenno sulle disequazioni variazionali in uno spazio di Banach riflessivo Σ, sul campo reale.

Detto K un sottoinsieme non vuoto di Σ, chiuso e convesso, siano:

$$A : \Sigma \to \Sigma^*$$

un operatore di Σ nel suo duale, non necessariamente lineare, ed f un elemento di Σ^*. L'operatore A può anche non essere definito in tutto Σ: basta che sia definito nell'insieme K.

Il problema di ricercare un elemento $u \in K$ tale che si abbia:

$$\langle Au, v-u \rangle \geq \langle f, v-u \rangle \quad \forall v \in K \tag{14}$$

è una generalizzazione di quello espresso dalla (4) di pag. 255, e si chiama anch'esso una <u>disequazione variazionale</u>. Il problema (4), con la forma bilineare $a(u,v)$ continua, si ottiene nel caso particolare che Σ sia uno spazio di Hilbert ed A sia l'operatore che ad ogni $u \in \Sigma$ associa il funzionale lineare e continuo $v \in \Sigma \to a(u,v)$, sicché in tal caso risulta:

$$\langle Au, v \rangle = a(u,v) .$$

Nel teorema di esistenza per la (14) l'operatore A non si suppone continuo. Si suppone però che per ogni coppia (u,v) di punti di K sia continua la funzione reale della variabile reale t definita ponendo:

$$\alpha(t) = \langle A[tu+(1-t)v], u-v \rangle \quad \forall t \in [0,1] .$$

Tale proprietà si esprime dicendo che l'operatore A è <u>emicontinuo</u>.

Nel caso che K non sia limitato si suppone poi che l'operatore A soddisfi alla seguente condizione: esiste un elemento v_0 di K tale che si abbia

$$\lim_{\substack{\|v\| \to +\infty \\ v \in K}} \frac{\langle Av, v-v_0 \rangle}{\|v\|} = +\infty .$$

Tale condizione si esprime dicendo che l'operatore A è <u>coercivo</u>.

Ciò premesso, sussistono i seguenti teoremi, di unicità e di esistenza:

(1.5) <u>Se l'operatore $A : K \to \Sigma^*$ è strettamente monotono (1), qualunque sia $f \in \Sigma^*$ la disequazione variazionale (14) ammette al più una soluzione</u>.

(1) Cfr. pag. 71. Ovviamente la definizione ivi introdotta ha senso anche se A non è definito in tutto Σ.

(1.6). Qualunque sia $f \in \Sigma^*$ la disequazione variazionale (14) ammette almeno una soluzione se K è limitato e l'operatore A è monotono col emicontinuo, oppure se K non è limitato e l'operatore A è monotono, emicontinuo e coercivo.

La (1.5) si dimostra con lo stesso procedimento seguito per la (1.1) di pag. 255, utilizzando la stretta monotonia di A in luogo della coercitività di $a(u,v)$. Della (1.6) non è il caso di riportare la dimostrazione.

Ovviamente nel caso che K sia un sottospazio di Σ, e perciò diciamolo H, la disequazione variazionale (14) equivale al problema di ricercare $u \in H$ tale che si abbia:

(15) $$\langle Au, z \rangle = \langle f, z \rangle \quad \forall z \in H \quad (1).$$

Tale problema, che generalizza quello espresso dalla (13) di pag. 257, si chiama una equazione variazionale. Dalle (1.5), (1.6) consegue:

(1.7) Qualunque sia $f \in \Sigma^*$, se l'operatore $A : H \to \Sigma^*$ è strettamente monotono l'equazione variazionale (15) ammette al più una soluzione, ed ammette almeno una soluzione se A è monotono, emicontinuo e coercivo.

Pur essendo ovvio osserviamo che se u è soluzione della disequazione variazionale (14), ed è un punto interno al convesso K, allora u è soluzione dell'equazione variazionale (15) con $H = \Sigma$.

Dim. Supponiamo invero che il cerchio aperto di Σ di centro u e raggio δ sia contenuto in K. Fissato un qualunque $z \in \Sigma$ e scelto un numero $\lambda > \|z\| / \delta$, il punto $\frac{1}{\lambda} z + u$ appartiene a detto cerchio, quindi a K. Ponendo nella (14) $v = \frac{1}{\lambda} z + u$, si ottiene $\langle Au, \frac{1}{\lambda} z \rangle \geq \langle f, \frac{1}{\lambda} z \rangle$, ossia $\langle Au, z \rangle \geq \langle f, z \rangle$. Poichè z è un qualunque elemento dello spazio, tale disuguaglianza sussiste anche con $-z$ in luogo di z. Allora vale anche la disuguaglianza opposta, e quindi vale la (15).

---

(1) Ciò si verifica con lo stesso ragionamento fatto nella nota (1) di pag. 257.

Concludiamo con la seguente osservazione (Lemma di Minty):

(1.8). Se l'operatore $A : K \to \Sigma^*$ è monotono ed emicontinuo, un elemento $u \in K$ è soluzione della disequazione variazionale:

$$(16) \qquad \langle Au, v-u \rangle \geq 0 \qquad \forall v \in K$$

se e solo se risulta:

$$(17) \qquad \langle Av, v-u \rangle \geq 0 \qquad \forall v \in K .$$

Dim. Essendo A monotono, qualunque sia $v \in K$ risulta:

$$\langle Av - Au, v-u \rangle \geq 0,$$

ossia:

$$\langle Av, v-u \rangle \geq \langle Au, v-u \rangle .$$

Di conseguenza se u verifica la (16), verifica anche la (17).

Viceversa, supponiamo che u soddisfi alla (17). Detto w un qualunque elemento di K, poniamo:

$$v = tw + (1-t)u = u + t(w-u) \qquad \forall t \in [0,1],$$

e utilizziamo la (17) per tale $v \in K$. Si ha:

$$\langle A[u+t(w-u)], t(w-u) \rangle \geq 0$$

e quindi, essendo $t \geq 0$:

$$\langle A[u+t(w-u)], w-u \rangle \geq 0 .$$

Per la supposta emicontinuità di A, passando al limite per $t \to 0$ si ottiene:

$$\langle Au, w-u \rangle \geq 0,$$

e poiché tale relazione sussiste per ogni $w \in K$, la (16) è dimostrata.

Ulteriori considerazioni sulle disequazioni variazionali saranno svolte nell'Appendice [cfr. pag. 312 e pag. 317].

ESEMPI

Nei numeri successivi avremo occasione di studiare alcune disequazioni ed alcune equazioni variazionali. Per ora vogliamo mostrare alcuni problemi che conducono a disequazioni variazionali.

1) Assegnata una funzione reale φ di classe C^1 in un intervallo $[a,b]$, determinare i punti di $[a,b]$ in cui φ assume il valore minimo.

Se un tale punto u è interno ad $[a,b]$, deve risultare $\varphi'(u) = 0$; se $u = a$,

deve essere $\varphi'(u) \geq 0$; se $u=b$, deve essere $\varphi'(u) \leq 0$. In ogni caso, deve risultare:

(18) $$\varphi'(u)(v-u) \geq 0 \quad \forall v \in [a,b].$$

Le soluzioni del problema vanno dunque ricercate fra i punti $u \in [a,b]$ che verificano la (18). Orbene la (18) è una disequazione variazionale del tipo (14) con $\Sigma = R$, $K=[a,b]$, $f=0$, essendo A l'operatore che ad $u \in K$ associa il funzionale $v \in R \to \varphi'(u)v$.

2) Assegnata una funzione reale $\varphi \in C^1(\bar{\Omega})$ con Ω aperto convesso di R^n, e posto $K = \bar{\Omega}$, determinare gli eventuali punti di K in cui φ assume il valore minimo.

Se $\varphi(u) = \min_K \varphi$, qualunque sia $v \in K$ la funzione definita ponendo:

$$F(t) = \varphi(tu + (1-t)v) \quad \forall t \in [0,1]$$

assume per $t=1$ il suo valore minimo, e quindi deve risultare $F'(1) \leq 0$. Essendo

$$F'(t) = \text{grad}\varphi(tu + (1-t)v) \times (u-v)$$

e quindi $F'(1) = \text{grad}\varphi(u) \times (u-v)$, deve risultare:

(19) $$(\text{grad}\varphi(u), v-u) \geq 0 \quad \forall v \in K.$$

La (19), che per $n=1$ e $K=[a,b]$ si riduce alla (18), è una disequazione variazionale del tipo (14) con $\Sigma = R^n$, $f=0$, essendo A l'operatore che ad $u \in K$ associa il funzionale:

$$v \in R^n \longrightarrow (\text{grad}\varphi(u), v).$$

3) Assegnate due funzioni reali $\psi_1, \psi_2 \in C^1([a,b])$, tra le funzioni $u \in C^1([a,b])$ nulle negli estremi di $[a,b]$ e tali che:

(20) $$\psi_1(x) \leq u(x) \leq \psi_2(x) \quad \forall x \in [a,b]$$

ricercare quelle (se ne esistono) che minimizzano l'integrale:

$$\int_a^b u'^2(x)\,dx.$$

Perché il problema abbia senso, ovviamente le funzioni ψ_1, ψ_2 debbono soddisfare alle condizioni: $\psi_1(a) \leq 0 \leq \psi_2(a)$, $\psi_1(b) \leq 0 \leq \psi_2(b)$.

Detto K l'insieme costituito dalle $u \in C^1([a,b])$ nulle negli estremi di $[a,b]$ e sodolisfacenti alla (20), ovviamente K è un sottoinsieme non vuoto e convesso di $C^1([a,b])$; utilizzando la (1.8) di pag. 109 e ricordando la convergenza di una successione nello spazio C^1 [pag. 101, prima dell'es. 6], è facile poi dimostrare che K è chiuso.

Se u e' una soluzione del problema, qualunque sia $v \in K$ la funzione definita ponendo:

$$F(t) = \int_a^b [t u'(x) + (1-t) v'(x)]^2 dx \qquad \forall t \in [0,1]$$

assume per $t=1$ il suo valore minimo, e quindi deve essere $F'(1) \leq 0$.

Essendo:

$$F'(t) = 2 \int_a^b [t u' + (1-t) v'] (u' - v') dx ,$$

deve dunque risultare:

$$\int_a^b u'(x) [v'(x) - u'(x)] dx \geq 0 \qquad \forall v \in K. \tag{21}$$

Se per ogni $u \in C^1([a,b])$ si denota con Au il funzionale lineare e continuo:

$$v \in C^1([a,b]) \longrightarrow \int_a^b u'(x) v'(x) dx ,$$

si definisce un operatore A di $C^1([a,b])$ nel suo duale $(C^1)^*$, e la (21) si scrive:

$$\langle Au, v-u \rangle \geq 0 \qquad \forall v \in K. \tag{22}$$

Il problema posto conduce dunque alla (22), che e' una disequazione variazionale del tipo (14).

Notiamo che la (21) ha significato anche per $u, v \in H_0^1(\Omega)$, con $\Omega =]a,b[$, naturalmente intendendo per K il sottoinsieme convesso di $H_0^1(\Omega)$ costituito dalle $u \in H_0^1(\Omega)$ che verificano la (20) quasi ovunque in (a,b) (1).

Generalizzando il problema in tal senso, e denotando con $(\cdot,\cdot)$ il prodotto scalare in $H_0^1(\Omega)$, cioe' ponendo:

$$(u,v) = \int_a^b u' v' dx \qquad \forall u, v \in H_0^1(\Omega),$$

per l'operatore A dianzi introdotto risulta:

$$\langle Au, v \rangle = (u,v)$$

e quindi la (22) diviene:

$$(u, v-u) \geq 0 \qquad \forall v \in K. \tag{23}$$

Dunque generalizzando il problema si ottiene una disequazione variazionale del tipo (4) di pag. 255.

---

(1) L'insieme K e' anche chiuso: rinunciamo alla dimostrazione di cio', anche perche' una questione dello stesso tipo sara' risolta nei numeri successivi.

4) Assegnate due funzioni reali $\psi_1, \psi_2 \in C^1(\bar{\Omega})$, con Ω aperto di R^n, tra le funzioni $u \in C^1(\bar{\Omega})$ nulle su $\partial\Omega$ e tali che:

$$\psi_1(x) \leq u(x) \leq \psi_2(x) \quad \forall x \in \bar{\Omega}, \tag{24}$$

ricercare quelle (se ne esistono) che minimizzano l'integrale:

$$\int_\Omega |\operatorname{grad} u|^2 dx .$$

Perché il problema abbia senso, per ogni $x \in \partial\Omega$ deve risultare ovviamente $\psi_1(x) \leq 0 \leq \psi_2(x)$.

Detto K l'insieme costituito dalle $u \in C^1(\bar{\Omega})$ nulle su $\partial\Omega$ e soddisfacenti alla (24), K è un sottoinsieme non vuoto di $C^1(\bar{\Omega})$, chiuso e convesso.

Con lo stesso ragionamento svolto nell'esempio precedente si conclude che, se u è una soluzione del problema, deve risultare:

$$\int_\Omega [\operatorname{grad} u \times \operatorname{grad}(v-u)] dx \geq 0 \quad \forall v \in K . \tag{25}$$

Se per ogni $u \in C^1(\bar{\Omega})$ si denota con Au il funzionale lineare e continuo:

$$v \in C^1(\bar{\Omega}) \to \int_\Omega [\operatorname{grad} u \times \operatorname{grad} v] dx ,$$

si definisce un operatore A di $C^1(\bar{\Omega})$ nel suo duale $(C^1)'$, e la (25) si scrive:

$$\langle Au, v-u \rangle \geq 0 \quad \forall v \in K .$$

Il problema posto conduce dunque ad una disequazione variazionale del tipo (14).

Nel caso $n=1$ il problema considerato si riduce a quello dell'esempio precedente. Come per l'esempio precedente, il problema si generalizza considerandolo nello spazio $H_0^1(\Omega)$ e conduce ad una disequazione variazionale del tipo (4) di pag. 255.

2. Formulazione variazionale di un particolare problema ai limiti.

Abbiamo già visto un esempio di equazione variazionale [es. 9 pag. 182 bis], come applicazione del teorema di Riesz. Con l'occasione abbiamo constatato che, per ogni funzionale f definito in $H_0^1(\Omega)$, lineare e continuo, l'equazione variazionale:

$$u \in H_0^1(\Omega): \qquad (u,v)_{H_0^1(\Omega)} = \langle f, v \rangle \quad \forall v \in H_0^1(\Omega)$$

esprime in senso generalizzato il problema omogeneo di Dirichlet per l'operatore di Laplace.

In questo numero consideriamo un problema ai limiti che ci occorrerà nel seguito, e dimostreremo in base al teorema di Lax-Milgram che esso ammette una e una sola soluzione.

In senso classico il problema è il seguente: assegnate due funzioni $c(x)$, $g(x)$ continue in un intervallo $[a,b]$, stabilire se esiste una funzione $u(x)$, di classe C^2 in $[a,b]$, che sia soluzione dell'equazione differenziale:

$$(1) \qquad -u'' + c(x)u = g(x)$$

e verifichi le condizioni ai limiti:

$$(2) \qquad u(a) = u(b) = 0 .$$

L'equazione (1) si può intendere nel senso di ricercare u tra le funzioni dotate di derivata seconda debole, in particolare nello spazio $H^2(\Omega)$, con $\Omega =]a,b[$ [1].

Se poi si ricerca u appartenente anche allo spazio $H_0^1(\Omega)$, ciò assicura [teor. di Sobolev, pag. 105] che esiste un rappresentante di u continuo in $\overline{\Omega}$ e nullo negli estremi di Ω, sicché tale rappresentante verifica le (2).

Dunque il problema si può generalizzare ricercando u nello spazio:

$$H = H^2(\Omega) \cap H_0^1(\Omega).$$

A tal fine non è necessario che i dati $c(x)$, $g(x)$ siano funzioni continue.

Esaminando il primo membro della (1), si vede che per $u \in H^2(\Omega)$ il primo addendo appartiene ad $L^2(\Omega)$, e affinché vi appartenga anche il secondo addendo è sufficiente che $c(x)$ sia limitata; dopo ciò, per avere senso l'uguaglianza è sufficiente che $g(x)$ appartenga ad $L^1(\Omega)$.

---

[1] È utile qui ricordare che se Ω è un intervallo di R lo spazio $W^2(\Omega)$ coincide con $H^2(\Omega)$ [cfr. es. 14 pag. 103].

Il problema considerato si può porre allora in senso generalizzato come segue: assegnati $c \in L^\infty(\Omega)$, $g \in L^2(\Omega)$, ricercare un elemento u dello spazio $H_0^1(\Omega)$, dotato di derivata seconda debole, tale che u'' coincida con l'elemento $cu-g$ di $L^2(\Omega)$.

Se esiste un siffatto elemento u, per definizione di derivata debole si ha:

$$(3) \qquad \int_\Omega (cu-g)\varphi\,dx = \int_\Omega u\varphi''dx \quad \forall \varphi \in C_0^\infty(\Omega),$$

ossia:

$$\int_\Omega (cu-g)\varphi\,dx = -\int_\Omega u'\varphi'dx \quad \forall \varphi \in C_0^\infty(\Omega) \quad (1).$$

Tale uguaglianza può scriversi:

$$\int_\Omega (u'\varphi' + cu\varphi)\,dx = \int_\Omega g\varphi\,dx \quad \forall \varphi \in C_0^\infty(\Omega),$$

ed equivale, essendo $C_0^\infty(\Omega)$ denso in $H_0^1(\Omega)$ [pag.104], alla seguente:

$$\int_\Omega (u'v' + cuv)\,dx = \int_\Omega gv\,dx \quad \forall v \in H_0^1(\Omega).$$

Se allora si pone, per $u, v \in H_0^1(\Omega)$:

$$a(u,v) = \int_\Omega (u'v' + cuv)\,dx \quad , \quad \langle f, v\rangle = \int_\Omega gv\,dx,$$

l'eventuale soluzione del problema è soluzione dell'equazione variazionale:

$$(4) \qquad u \in H_0^1(\Omega): \quad a(u,v) = \langle f, v\rangle \quad \forall v \in H_0^1(\Omega).$$

Viceversa se esiste una soluzione u della (4), ossia della (3), in virtù della (3) l'elemento $cu-g$ di $L^2(\Omega)$ è la derivata seconda debole di u, e quindi u è soluzione del problema generalizzato.

In definitiva il problema in senso generalizzato è equivalente all'equazione variazionale (4). Orbene, il funzionale f è continuo [es.10 pag.148]; la forma bilineare $a(u,v)$ è continua, ed è coercitiva se $c \geq 0$ [es.1 pag.213]: pertanto, per il teorema di Lax-Milgram, se $c \geq 0$ la (4) ammette una e una sola soluzione.

Riassumiamo la conclusione con il seguente enunciato:

---

(1) Per definizione di derivata debole risulta:

$$\int_\Omega u'\psi\,dx = -\int_\Omega u\psi'dx \quad \forall \psi \in C_0^\infty(\Omega)$$

da cui, per $\psi = \varphi'$, consegue:

$$\int_\Omega u'\varphi'\,dx = -\int_\Omega u\varphi''dx \quad \forall \varphi \in C_0^\infty(\Omega).$$

(2.1). Qualunque siano $g \in L^2(\Omega)$ e $c \in L^\infty(\Omega)$ con $c \geq 0$, l'equazione:

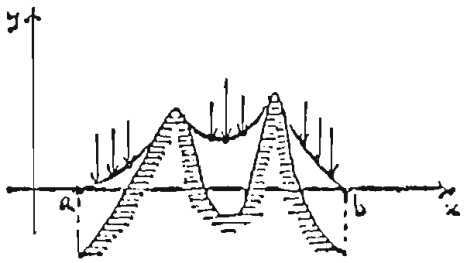
(4) $$-u'' + c(x)u = g$$

ammette nello spazio $H^2(\Omega) \cap H_0^1(\Omega)$ una e una sola soluzione.

3. Un problema sul filo flessibile teso su un ostacolo rigido.

Consideriamo il seguente problema: determinare la configurazione di equilibrio di un filo flessibile fissato agli estremi, teso al disopra di un ostacolo rigido e soggetto ad un carico distribuito (forza posizionale distribuita sul filo) rappresentato da un sistema di vettori piano. Il contatto filo-ostacolo si suppone privo di attrito.

Studiare il problema significa in primo luogo tradurlo in termini matematici, ossia formulare un problema puramente matematico che possa rappresentare efficacemente un "modello" della questione fisica proposta.



Supponiamo dunque che gli estremi del filo siano fissati in due punti a e b dell'asse x, e che l'ostacolo sia rappresentato dal diagramma di una funzione $\psi(x)$ continua nell'intervallo $[a,b]$, con $\psi(a) < 0$, $\psi(b) < 0$; rappresentiamo quindi la configurazione di equilibrio del filo mediante la curva di equazione $y = u(x)$.

La funzione u deve essere nulla negli estremi dell'intervallo $[a,b]$, e deve soddisfare l'equazione della funicolare:

(1) $$-\tau u'' = F,$$

dove τ è la componente orizzontale della tensione del filo, e F è la forza distribuita risultante delle forze agenti sul filo.

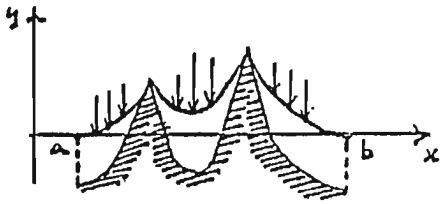
Esprimiamo mediante una funzione continua f il carico distribuito sul filo, e denotiamo con f_1 la funzione che rappresenta le forze dovute alla presenza dell'ostacolo, che hanno la direzione e il verso dell'asse y: si ha dunque $F = f + f_1$, con $f_1 \geq 0$. Per evitare inutili complicazioni formali supponiamo $\tau = 1$, la (1) fornisce allora $-u'' = f + f_1$, e quindi $-u'' \geq f$.

Naturalmente deve essere $u \geq \psi$; poichè nei punti in cui il filo non tocca l'ostacolo l'unica forza agente sul filo è f, sicchè $F = f$, per ogni x tale che $u(x) > \psi(x)$ deve poi risultare $-u''(x) = f(x)$.

In definitiva la funzione $u(x)$ deve soddisfare alle seguenti condizioni:

$$(2)\qquad \begin{cases} -u'' \geq f \\ u(a)=u(b)=0 \\ u \geq \psi \\ u(x)>\psi(x) \Longrightarrow -u''(x)=f(x) \end{cases}$$

Nasce così il problema di ricercare una funzione $u(x)$, di classe $C^{(2)}$ in $[a,b]$, soddisfacente alle (2).



Ci si rende subito conto, però, che la configurazione di equilibrio del filo non sempre può essere rappresentata da una funzione di classe $C^{(2)}$: basta osservare la figura qui accanto. E' ben vero che nella realtà non esistono ostacoli "a punta", e che pertanto nella realtà la funzione $u(x)$ ha derivata continua; peraltro i modelli matematici devono contemplare anche i casi limiti, e questi, pur non rispecchiando fedelmente la realtà, conducono poi a conclusioni ben aderenti ai fenomeni reali: si pensi ad esempio ai "punti materiali", alla "velocità istantanea"; agli urti, in cui il vettore velocità si muta nel vettore opposto "istantaneamente".

Dunque per tradurre efficacemente in termini matematici il problema fisico considerato, bisogna intendere le (2) in un senso più generale.

Se si ricerca u nello spazio $H_0^1(\Omega)$, con $\Omega=]a,b[$, la seconda delle (2) sarà verificata automaticamente dal rappresentante di u continuo in $[a,b]$ e nullo negli estremi [teor. di Sobolev, pag. 105]; la condizione $u \geq \psi$ si dovrà intendere sodisfatta "quasi ovunque", e le rimanenti condizioni si possono intendere nel senso delle distribuzioni.

Formuliamo allora il problema in senso generalizzato come segue:

Assegnati $f \in L^2(\Omega)$, $\psi \in C^0(\bar{\Omega})$ con $\psi(a)<0$, $\psi(b)<0$, ricercare $u \in H_0^1(\Omega)$ tale che si abbia:

$$(3)\qquad \begin{cases} -u'' \geq f & \text{nel senso delle distribuzioni su } \Omega \\ u \geq \psi & \text{quasi ovunque in } \Omega \\ -u'' = f & \text{nel senso delle distribuzioni su } \Omega' \,. \end{cases}$$

dove Ω' è l'insieme costituito dai punti $x \in \Omega$ tali che $\tilde{u}(x)>\psi(x)$, essendo $\tilde{u}$ il rappresentante continuo di u.

A questo punto occorre stabilire se le condizioni imposte alla u sono sufficienti a determinarla univocamente: ciò assicura che non sono state omesse delle considerazioni fisiche, che avrebbero condotto ad ulteriori condizioni. Ma bisogna anche accertare che le condizioni imposte non sono troppo restrittive, come accadeva per il problema (2): ciò si garantisce dimostrando che il problema è risolubile, ossia che ammette soluzioni.

In breve: occorre un teorema di esistenza e unicità. Dopo ciò, detta Φ la corrispondenza che ad ogni $f \in L^2(\Omega)$ associa la soluzione u del problema, per concludere che esso è ben posto occorre dimostrare che la funzione Φ è continua. Non ci soffermeremo su quest'ultima affermazione, limitando il nostro studio all'esistenza e unicità: ciò vale anche per gli altri problemi che presenteremo nei numeri successivi.

A proposito della locuzione "problema ben posto", consigliamo al lettore di ritornare su quanto detto a pag. 182.5.

* * *

Vogliamo ora mostrare che effettivamente il problema (3) ammette una e una sola soluzione. A ciò si perviene esprimendo il problema mediante una disequazione variazionale.

Consideriamo intanto il sottoinsieme di $H_0^1(\Omega)$:

$$K = \{ v \in H_0^1(\Omega) : v \geq \psi \text{ q.o.} \} \tag{4}$$

e dimostriamo che esso è non vuoto, chiuso e convesso.

Dim. Per evitare inutili complicazioni formali, supponiamo che l'insieme $\Omega =]a,b[$ sia l'intervallo $]-l,l[$. Scegliamo due numeri positivi α e δ in modo che si abbia:

$$\psi(x) < 0 \text{ in } \overline{\Omega} - [-\alpha,\alpha], \quad \alpha < \delta, \quad [-\delta,\delta] \subset \Omega,$$

nonché un numero h tale da risultare:

$$h > \frac{\max|\psi|}{(\delta^2-\alpha^2)^2}.$$

La funzione:

$$v(x) = \begin{cases} 0 & \text{se } x \in \overline{\Omega} - [-\delta,\delta] \\ h(\delta^2-x^2)^2 & \text{se } x \in [-\delta,\delta] \end{cases}$$

è un elemento di $H_0^1(\Omega)$, in quanto appartiene a $C_0^2(\Omega)$. Per $|x| > \alpha$ si ha poi

$v(x) \geq 0 > \psi(x)$, e per $|x| \leq \alpha$ si ha (essendo di conseguenza $|x| < \delta$):

$$v(x) = h(\delta^2 - x^2)^2 \geq h(\delta^2 - \alpha^2)^2 > \max|\psi| \geq \psi(x).$$

Dunque risulta $v(x) > \psi(x)$ $\forall x \in \Omega$, sicché $v \in K$. Pertanto K è non vuoto.

Che K sia convesso è ovvio. Per dimostrare che K è chiuso, utilizziamo il seguente risultato: <u>se $\varphi \in L^2(\Omega)$ e se $\{\varphi_m\}$ è una successione di punti di $L^2(\Omega)$ tendente a φ in $L^2(\Omega)$, dalla $\{\varphi_m\}$ si può estrarre una successione che converge a φ quasi ovunque in Ω.</u>

Detta allora $\{v_m\}$ una successione di punti di K convergente a v_0 in $H_0^1(\Omega)$, esistono una estratta $\{v_{k_m}\}$ e un sottoinsieme Ω_0 di Ω di misura nulla tali che si abbia:

$$\lim v_{k_m}(x) = v_0(x), \quad v_{k_m}(x) \geq \psi(x) \qquad \forall x \in \Omega - \Omega_0.$$

Pertanto risulta $v_0(x) \geq \psi(x)$ $\forall x \in \Omega - \Omega_0$, e quindi $v_0 \in K$. Ne consegue che K è chiuso.

— — — — — — — — —

Identificando poi l'elemento f di $L^2(\Omega)$ con il funzionale lineare e continuo [es.10 pag.148]:

$$v \in H_0^1(\Omega) \longrightarrow \int_\Omega f v \, dx,$$

dimostriamo che:

(3.1) <u>Il problema di ricercare un elemento u di $H_0^1(\Omega)$ soddisfacente alle (3) equivale a quello di ricercare una soluzione u della disequazione variazionale:</u>

$$(5) \qquad u \in K: \quad (u, v-u)_{H_0^1(\Omega)} \geq \langle f, v-u \rangle \quad \forall v \in K.$$

— — — — — — — — —

Dim. (3)$\Rightarrow$(5). Supponiamo che l'elemento u di $H_0^1(\Omega)$ soddisfi alle (3).

Per la prima delle (3) la distribuzione $-(u''+f)$ è positiva [1], e quindi esiste una misura μ, definita nell'insieme delle parti misurabili di Ω, che assume valore finito sui compatti, ed è tale da risultare [cfr.(18) pag.152]:

$$(6) \qquad \langle -(u''+f), \varphi \rangle = \int_\Omega \varphi \, d\mu \qquad \forall \varphi \in C_0^\infty(\Omega)$$

ossia, per ogni $\varphi \in C_0^\infty(\Omega)$:

$$\langle -u'', \varphi \rangle - \langle f, \varphi \rangle = \int_\Omega \varphi \, d\mu$$

cioè [2]:

$$\int_\Omega u' \varphi' \, dx - \langle f, \varphi \rangle = \int_\Omega \varphi \, d\mu$$

cioè ancora [3]:

$$(u, \varphi)_{H_0^1(\Omega)} - \langle f, \varphi \rangle = \int_\Omega \varphi \, d\mu.$$

Fissato $v \in K$, e detta $\{\varphi_m\}$ una successione di elementi di $C_0^\infty(\Omega)$ convergente a $v-u$ in $H_0^1(\Omega)$ [4], si ha allora:

---

(1) Ricordiamo [pag.156] che la distribuzione u'' è il funzionale $\varphi \in C_0^\infty(\Omega) \to \int_\Omega u\varphi'' \, dx$, e che f si identifica con la distribuzione che ad ogni $\varphi \in C_0^\infty(\Omega)$ associa l'integrale $\int_\Omega f\varphi$.

(2) Cfr. (1), tenendo conto di quanto detto nella nota (1) di pag.259.

(3) Si ricordi il prodotto scalare in $H_0^1(\Omega)$ [pag.104].

(4) $C_0^\infty(\Omega)$ è denso in $H_0^1(\Omega)$ [pag.104].

$$(7)\qquad (u,\varphi_n)_{H_0^1(\Omega)} - \langle f,\varphi_n\rangle = \int_\Omega \varphi_n\, d\mu .$$

Per $n\to\infty$ il primo membro della (7) tende a

$$(u,v-u)_{H_0^1(\Omega)} - \langle f, v-u\rangle$$

e perciò in virtù della (7), per dimostrare che u è soluzione della (5), si tratta di far vedere che si ha:

$$(8)\qquad \lim \int_\Omega \varphi_n\, d\mu \geq 0 .$$

Osserviamo intanto che per la terza delle (3), per tutte le φ che appartengono a $C_0^\infty(\Omega')$ [1] il primo membro della (6) è nullo, e l'integrale a secondo membro si può estendere ad Ω' (in quanto $\varphi=0$ in $\Omega-\Omega'$). Per la (6) stessa si ha allora:

$$\int_{\Omega'} \varphi\, d\mu = 0 \quad \forall \varphi \in C_0^\infty(\Omega'),$$

e quindi risulta [1]:

$$(9)\qquad \mu(\Omega') = 0$$

D'altra parte, detto $\tilde u$ il rappresentante continuo di u, siccome la funzione $\tilde u - \psi$ assume valore positivo nei punti a e b, esiste $\delta>0$ tale che in ciascuno degli intervalli $\Omega_1 =]a, a+\delta[$, $\Omega_2 =]b-\delta, b[$ si abbia $\tilde u - \psi > 0$, sicché $\Omega_1 \subseteq \Omega'$, $\Omega_2 \subseteq \Omega'$. Per la (9) si ha quindi $\mu(\Omega_1) = \mu(\Omega_2) = 0$; se allora si considera il compatto $C = [a+\frac{\delta}{2}, b-\frac{\delta}{2}]$, risultando $\Omega = \Omega_1 \cup \Omega_2 \cup C$ si ha:

$$\mu(\Omega) \leq \mu(\Omega_1) + \mu(\Omega_2) + \mu(C) = \mu(C) < +\infty$$

Dunque risulta:

$$(10)\qquad \mu(\Omega) < +\infty .$$

Osserviamo ora che, detto $\tilde v$ il rappresentante continuo di v, si ha:

$$(11)\qquad \tilde v \geq \psi \quad , \quad \psi = \tilde u \qquad \mu\text{-quasi ovunque in } \Omega,$$

la prima disuguaglianza perché valida ovunque in Ω, la seconda perché l'insieme dei punti di Ω in cui $\psi \neq \tilde u$, ossia $\psi < \tilde u$ per la seconda delle (3), è l'insieme Ω' che ha misura μ nulla [cfr. (9)]. Dalle (11) consegue $\tilde v - \tilde u = \tilde v - \psi \geq 0$ μ-quasi ovunque in Ω, e quindi:

$$(12)\qquad \int_\Omega (\tilde v - \tilde u)\, d\mu \geq 0 .$$

Siccome per il teorema di Sobolev risulta [cfr. (5) pag. 105]:

$$\| (\tilde v - \tilde u) - \varphi_n \|_{C^{0,\frac{1}{2}}(\bar\Omega)} \leq c \, \| (v-u) - \varphi_n \|_{H_0^1(\Omega)} \qquad \forall n \in N$$

con c costante positiva, si ha:

$$(13)\qquad \lim \varphi_n = \tilde v - \tilde u \quad \text{uniformemente in } \bar\Omega$$

e quindi, tenuto conto della (10) e utilizzando poi la (12):

$$\lim \int_\Omega \varphi_n\, d\mu = \int_\Omega (\tilde v - \tilde u)\, d\mu \geq 0 \quad [3].$$

---

[1] L'insieme Ω' è stato introdotto a pag. 261 per formulare le (3).

[2] Se $\mu(\Omega') \neq 0$, è possibile costruire una $\varphi \in C_0^\infty(\Omega')$ il cui integrale rispetto a μ esteso ad Ω' sia non nullo.

[3] Per le (10) e (13) è lecito il passaggio al limite sotto il segno di integrale.

La (8) è pertanto acquisita, e con ciò l'asserto è dimostrato [1].

— — — — — — — — —

Dim. (5) $\Rightarrow$ (3). Supponiamo che l'elemento u di $H_0^1(\Omega)$ sia soluzione della (5).

Qualunque sia $\varphi \in C_0^\infty(\Omega)$ tale che

(14) $$\varphi \geqslant 0 \ ,$$

l'elemento $v = u + \varphi$ di $H_0^1(\Omega)$ appartiene a K. Ponendo allora nella (5) $v = u + \varphi$ si ottiene:

$$\int_\Omega u' \varphi' \, dx \geqslant \langle f, \varphi \rangle$$

cioè [cfr. nota (4) pag. 263]:

$$\langle -u'', \varphi \rangle \geqslant \langle f, \varphi \rangle$$

ossia:

$$\langle -(u'' + f), \varphi \rangle \geqslant 0 .$$

Questa relazione, soddisfatta $\forall \varphi \in C_0^\infty(\Omega)$ tale che valga la (14), significa che la distribuzione $-(u'' + f)$ è positiva, cioè che $-u'' \geqslant f$ nel senso delle distribuzioni. Pertanto u verifica la prima delle (3).

Per provare la terza delle (3), si tratta di dimostrare che risulta:

$$-\int_{\Omega'} u \varphi'' \, dx = \int_{\Omega'} f \varphi \, dx \qquad \forall \varphi \in C_0^\infty(\Omega')$$

ossia, tenuto conto che le $\varphi \in C_0^\infty(\Omega')$ sono nulle in $\Omega - \Omega'$:

$$-\int_\Omega u \varphi'' \, dx = \int_\Omega f \varphi \, dx \qquad \forall \varphi \in C_0^\infty(\Omega').$$

ossia ancora [cfr. (1) pag. 259]:

(15) $$\int_\Omega u' \varphi' \, dx = \int_\Omega f \varphi \, dx \qquad \forall \varphi \in C_0^\infty(\Omega').$$

Fissiamo allora $\varphi \in C_0^\infty(\Omega')$, non identicamente nulla, e denotiamo con $\tilde{u}$ il rappresentante continuo di u.

Detto S il supporto di φ, essendo S compatto la funzione continua $\tilde{u} - \psi$ è ivi dotata di minimo, diciamolo m, e siccome $\tilde{u} - \psi$ è positiva in quanto $S \subset \Omega'$, risulta $m > 0$. Posto $M = \max_S |\varphi|$, sia ε un numero positivo minore di m/M. Si ha:

$$\varepsilon < \frac{m}{M} \leqslant \frac{\tilde{u}(x) - \psi(x)}{M} \qquad \forall x \in S$$

e quindi:

$$\varepsilon M < \tilde{u}(x) - \psi(x) \qquad \forall x \in S \ ,$$

sicché risulta:

$$\varepsilon \varphi(x) \leqslant \varepsilon |\varphi(x)| \leqslant \varepsilon M < \tilde{u}(x) - \psi(x) \qquad \forall x \in S$$

e quindi, essendo $\varphi = 0$ in $\Omega - S$:

$$\tilde{u}(x) - \varepsilon \varphi(x) \geqslant \psi(x) \qquad \forall x \in \Omega .$$

Pertanto $u - \varepsilon\varphi$ è un elemento del convesso K. Ponendo allora in (5) $v = u - \varepsilon\varphi$, si ottiene:

---

(1) Tenuto conto che $u \in K$ per la seconda delle (3).

$$-\varepsilon \int_\Omega u'\varphi' dx \geq -\varepsilon \int_\Omega f\varphi\, dx ,$$

da cui:

$$\int_\Omega u'\varphi' dx \leq \int_\Omega f\varphi\, dx .$$

Poichè tale relazione deve valere $\forall \varphi \in C_0^\infty(\Omega)$, essa deve sussistere anche se in essa si sostituisce φ con $-\varphi$. Ne segue che vale anche la disuguaglianza opposta, e pertanto la (15) è dimostrata.

Poichè la seconda delle (3) è soddisfatta in quanto u appartiene a K, l'assunto è provato.

— — — — — — — — — —

In virtù del teorema di Stampacchia [(1.1) pag.255], dalla (3.1) consegue che (*):

(3.2) <u>Qualunque siano $f \in L^2(\Omega)$ e $\psi \in C^0(\bar\Omega)$ con $\psi(a)<0$, $\psi(b)<0$, esiste uno ed un solo elemento u di $H_0^1(\Omega)$ tale da verificare le (3).</u>

La soluzione u del problema, essendo elemento di $H_0^1(\Omega)$, è una classe di funzioni definite quasi ovunque in Ω, ed è chiaro che una tale classe non è atta a rappresentare la configurazione di equilibrio del filo. Quest'ultima è ovviamente il diagramma del rappresentante $\hat u$ continuo, esistente per il teorema di Sobolev [pag.105], della classe u.

Notiamo che per la (1.2) di pag.256, dal teor. (3.1) consegue che la soluzione u del problema è il punto che minimizza nel convesso K il funzionale:

$$\Phi(v) = \frac{1}{2}\int_\Omega v'^2 dx - \int_\Omega f v\, dx .$$

Pertanto fra tutti gli elementi v di $H_0^1(\Omega)$ tali che $v \geq \psi$ in Ω, l'elemento u è quello in cui il funzionale Φ assume il valore minimo. Da notare che il diagramma del rappresentante continuo di ciascuno dei predetti v si interpreta come una delle configurazioni possibili del filo.

Vogliamo aggiungere ancora un'osservazione. Essendo f un funzionale lineare e continuo in $H_0^1(\Omega)$, per il teorema di Riesz [(7.1) pag.177] esiste uno ed un solo punto $w \in H_0^1(\Omega)$ tale da risultare:

$$\langle f, v \rangle = (w, v) \quad \forall v \in H_0^1(\Omega).$$

Esprimendo mediante w il secondo membro della disuguaglianza variazionale (5) [pag.263] e ricordando la (4.2) di pag.117, si deduce che la soluzione u del problema è la proiezione di w sul convesso K.

(*) È banale che il prodotto scalare è una forma bilineare coercitiva.

4. Un problema sul filo flessibile teso su un ostacolo deformabile elasticamente.

Consideriamo il seguente problema: determinare la configurazione di equilibrio di un filo flessibile fissato agli estremi, teso al disopra di un ostacolo deformabile elasticamente, e soggetto ad un carico distribuito. Il contatto filo-ostacolo si suppone privo di attrito.

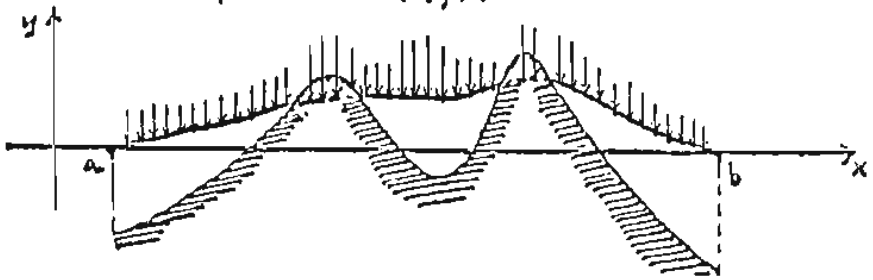
Supponiamo che gli estremi del filo siano fissati in due punti a e b dell'asse x, e che l'ostacolo sia rappresentato dal diagramma di una funzione $\psi(x)$ continua nell'intervallo $[a,b]$, con $\psi(a)<0$, $\psi(b)<0$; rappresentiamo quindi la configurazione di equilibrio del filo mediante la curva di equazione $y=u(x)$.

Supponiamo poi che il carico distribuito sul filo (sistema di vettori piano) sia rappresentato da una funzione $f\in L^2(\Omega)$, con $\Omega=]a,b[$, e che la deformazione dell'ostacolo, rappresentato da una funzione ψ continua in $[a,b]$ e tale che $\psi(a)<0$, $\psi(b)<0$, sia proporzionale alla reazione vincolare, secondo un coefficiente di proporzionalità (eventualmente variabile da punto a punto) espresso da una funzione $E\in L^\infty(\Omega)$ essenzialmente positiva in senso stretto, ossia tale da aversi:

$$E(x)\geq c_0>0 \quad \text{q.o. in } \Omega \qquad (c_0=\text{costante}). \tag{1}$$

Nella posizione di equilibrio, il fatto che il filo in un punto (di ascissa) x abbia deformato l'ostacolo sarà naturalmente rappresentato dalla disuguaglianza $u(x)<\psi(x)$, sicché la differenza $\psi(x)-u(x)$ esprimerà la deformazione dell'ostacolo in detto punto. Se $\varphi(x)$ è la reazione vincolare, sarà quindi:

$$\psi(x)-u(x)=E(x)\varphi(x). \tag{2}$$



Detta F la forza distribuita risultante delle forze agenti sul filo, e detta τ la componente orizzontale della tensione del filo, deve risultare:

$$-\tau u''=F \tag{3}$$

con la condizione:

$$u(a)=u(b)=0. \tag{4}$$

Nei punti x in cui il filo non tocca l'ostacolo, ossia tali che $u(x)>\psi(x)$, si ha $F=f$; nei punti x in cui il filo tocca (ed eventualmente deforma) l'ostacolo, ossia tali che $u(x)\leq\psi(x)$, si ha $F=f+\varphi$ e quindi, per la (2):

$$F=f+\frac{\psi-u}{E}.$$

Pertanto, supposto per semplicità $\tau = 1$, la (3) è espressa dalle condizioni:

$$u(x) > \psi(x) \Rightarrow -u''(x) = f(x) \tag{5}$$

$$u(x) \leq \psi(x) \Rightarrow -u''(x) = f(x) + \frac{\psi(x) - u(x)}{E(x)} \,. \tag{6}$$

Cercheremo la soluzione nello spazio:

$$H = H^2(\Omega) \cap H_0^1(\Omega) \tag{7}$$

che è uno spazio di Hilbert col prodotto scalare:

$$(u,v)_H = \int_\Omega (uv + u'v' + u''v'') \, dx \tag{8}$$

Per il teorema di Sobolev [pag. 105] gli elementi di H hanno un rappresentante di classe $C^1[\bar{\Omega}]$ che si annulla negli estremi di Ω, sicché tali elementi soddisfano automaticamente alla (4) [anzi sono che la verificano i suddetti rappresentanti].

Quanto alle (5) e (6), in cui figura la derivata seconda debole di u, sono da intendersi come segue. Fissato un rappresentante di u (che denotiamo ancora con u), e posto:

$$\Omega_1 = \{x \in \Omega : u(x) > \psi(x)\} \quad , \quad \Omega_2 = \{x \in \Omega : u(x) \leq \psi(x)\} \tag{9}$$

deve risultare:

$$-u'' = f \ \text{ q.o. in } \Omega_1 \ , \quad -u'' = f + \frac{\psi - u}{E} \ \text{ q.o. in } \Omega_2 .$$

Formuliamo dunque il seguente

PROBLEMA (4.1).

Assegnati $f \in L^2(\Omega)$ e $\psi \in C^0(\bar{\Omega})$, con $\psi(a) < 0$, $\psi(b) < 0$, si cercano $u \in H$ tale che si abbia:

$$\begin{cases} -u'' = f & \text{q.o. in } \Omega_1 \\ -u'' = f + \dfrac{\psi - u}{E} & \text{q.o. in } \Omega_2 \end{cases} \tag{10}$$

essendo H lo spazio (7) e Ω_1, Ω_2 gli insiemi (9).

Ci proponiamo di stabilire per tale problema un teorema di esistenza e unicità. A ciò si perviene esprimendo il problema mediante una disequazione variazionale, come si è fatto per quello del numero precedente.

E' opportuno però, preliminarmente, esprimere le condizioni (10) in un'altra forma nella quale non figurino i sottoinsiemi Ω_1 ed Ω_2.

Dimostriamo dunque in primo luogo che:

(4.2) Un elemento u dello spazio H soddisfa alle (10) se e solo se verifica quasi ovunque in Ω le relazioni:

$$u - Eu'' \geq \psi + Ef \tag{11}$$

$$u'' + f \leq 0 \tag{12}$$

$$(u''+f)(u - Eu'' - \psi - Ef) = 0 . \tag{13}$$

Dim. Supponiamo che u soddisfi alle (10). Fissato un rappresentante di u, proveremo che u verifica le (11), (12), (13) quasi ovunque in Ω_1 e quasi ovunque in Ω_2: di conseguenza le verificheranno quasi ovunque in $\Omega = \Omega_1 \cup \Omega_2$.

Nell'insieme Ω_1 per la prima delle (10) sono verificate banalmente la (12) e la (13), quasi ovunque; inoltre, essendo $E(u''+f)=0$ e $u-\psi>0$, risulta quasi ovunque:

$$u - \psi > E(u''+f),$$

ossia è verificata q.o. la (11).

Nell'insieme Ω_2 per la seconda delle (10) risulta, quasi ovunque:

$$u - Eu'' - \psi - Ef = 0,$$

e quindi sussistono q.o. la (11) e la (13); inoltre da $\psi - u \geq 0$, essendo per la seconda delle (10):

$$-u'' = f + (\psi - u)\frac{1}{\varepsilon},$$

segue $-u'' \geq f$ q.o., e quindi sussiste anche la (12).

Viceversa, supponiamo che u verifichi le (11), (12), (13) q.o. in Ω.

Fissato un rappresentante di ciascuna delle funzioni u, u'', E, f, denotiamo con Ω_0 l'insieme dei punti di Ω in cui le (11), (12), (13) sono verificate ovunque da tali rappresentanti (sicché l'insieme $\Omega - \Omega_0$ ha misura nulla). Proveremo che la prima [la seconda] delle (10) sussiste in ogni punto di $\Omega_1 \cap \Omega_0$ [di $\Omega_2 \cap \Omega_0$], sicché essa sussisterà quasi ovunque in Ω_1 [in Ω_2].

Sia $x \in \Omega_1 \cap \Omega_0$. Per dimostrare che in x risulta $u'' + f = 0$ basta far vedere, per la (13), che in x non può risultare:

$$u - Eu'' > \psi + Ef,$$

ossia:

$$u - \psi > E(u''+f). \tag{14}$$

Ed infatti, se nel punto x sussistesse la (14), per la (12) sarebbe $u(x) - \psi(x) \leq 0$, il che è assurdo in quanto x appartiene ad Ω_1.

Sia $x \in \Omega_2 \cap \Omega_0$. Per dimostrare che in x sussiste la seconda delle (10) basta far vedere, per la (11), che in x non può risultare:

$$u - Eu'' > \psi + Ef,$$

ossia

$$u - \psi > E(u''+f). \tag{15}$$

Ed infatti, se nel punto x sussistesse la (15), per la (13) in tale punto dovrebbe annullarsi $u''+f$, e dalla (15) stessa seguirebbe $u(x) - \psi(x) > 0$, il che è assurdo in quanto x appartiene ad Ω_2.

Consideriamo ora l'insieme:

$$K = \{ v \in H ;\ v - Ev'' \geq \psi + Ef \ \text{ q.o. in } \Omega \} \tag{16}$$

ed osserviamo che esso è non vuoto, chiuso e convesso.

Dim. Applichiamo la (2.1) di pag. 260 con $c = \frac{1}{E}$, $g = f + \frac{\psi}{E}$, il che è lecito in virtù della (1) di pag. 267 e delle ipotesi poste nel problema (4.1) su f e ψ. Se ne deduce che esiste $u \in H$ soluzione dell'equazione:

$$-u'' + \frac{1}{E} u = f + \frac{\psi}{E} ,$$

sicché tale soluzione appartiene a K. Pertanto K è non vuoto.

Che K sia convesso è ovvio. Per dimostrare che K è chiuso, sia $\{v_n\}$ una successione di elementi di K convergente a v_0 nello spazio H. Si ha allora:

$$v_n \to v_0 \ \text{ in } L^2(\Omega), \quad v_n' \to v_0' \ \text{ in } L^2(\Omega), \quad v_n'' \to v_0'' \ \text{ in } L^2(\Omega). \tag{17}$$

Denotiamo con (T) il teorema enunciato a pag. 263 per dimostrare che è chiuso l'insieme (4) di pag. 262. Dalla successione $\{v_n\}$ si può estrarre, in virtù di (T), una successione convergente a v_0 quasi ovunque in Ω, diciamola $\{w_{k_n}\}$; siccome per la seconda delle (17) si ha $w_n' \to v_0'$ in $L^2(\Omega)$, ancora per (T) dalla $\{w_n'\}$ si può estrarre una successione $\{w_{k_n}'\}$ convergente a v_0' quasi ovunque in Ω. Si ha allora:

$$w_{k_n} \to v_0 \ \text{ q.o. in } \Omega, \quad w_{k_n}' \to v_0' \ \text{ q.o. in } \Omega . \tag{18}$$

Siccome la successione $\{w_{k_n}''\}$ è estratta dalla $\{v_n''\}$ e quindi, per la terza delle (17), converge a v_0'' in $L^2(\Omega)$, dalla $\{w_{k_n}''\}$ si può estrarre, in virtù di (T), una successione convergente q.o. in Ω a v_0'', diciamola $\{w_{k_{z_n}}''\}$. Posto per semplicità di scrittura $w_{k_{z_n}} = z_n$, si ha allora, tenuto conto delle (18):

$$z_n \to v_0 , \quad z_n' \to v_0' , \quad z_n'' \to v_0'' \quad \text{q.o. in } \Omega ,$$

essendo $\{z_n\}$ una successione di punti dell'insieme (16). Allora dalla relazione:

$$\forall n \in N, \qquad z_n - E z_n'' \geq \psi + Ef \quad \text{q.o. in } \Omega$$

si deduce in modo ovvio che risulta:

$$v_0 - E v_0'' \geq \psi + Ef \quad \text{q.o. in } \Omega$$

e quindi $v_0 \in K$. Pertanto K è chiuso.

Orbene, posto:

$$a(u,v) = \int_\Omega (u'v' + Eu''v'')\, dx \qquad \forall u, v \in H \tag{19}$$

$$\langle F, v \rangle = \int_\Omega f \cdot (v - Ev'')\, dx \qquad \forall v \in H , \tag{20}$$

siamo ora in grado di dimostrare che:

(4.3). Il problema (4.1) è equivalente alla disequazione variazionale:

(21) $\qquad u \in K : \quad a(u, v-u) \geq \langle F, v-u \rangle \quad \forall v \in K.$

Dim. (4.1) $\Rightarrow$ (21). Supponiamo che u sia soluzione del problema (4.1), ossia, per la (4.2) di pag. 269, che u soddisfi alle (11), (12), (13).

Per la (11), u appartiene a K.

Sia v un elemento di K, cioè un elemento di H tale che:

$$-\psi \geq -v + Ev'' + Ef \qquad \text{q.o. in } \Omega.$$

Utilizzando questa disuguaglianza, minoriamo il secondo dei due fattori a primo membro della (13). Si ha:

$$u - Eu'' - \psi - Ef \geq u - Eu'' + (-v + Ev'' + Ef) - Ef = (u-v) - E(u''-v'')$$

da cui, moltiplicando per $(-u''-f)$ che è ≥ 0 per la (12), si deduce:

$$-(u''+f)(u - Eu'' - \psi - Ef) \geq -u''(u-v) - f(u-v) + Eu''(u''-v'') + Ef(u''-v'') =$$

$$= -u''(u-v) + Eu''(u''-v'') - f[(u-v) - E(u''-v'')].$$

Siccome per la (13) il primo membro è nullo, si ha allora:

$$u''(u-v) - Eu''(u''-v'') + f[(u-v) - E(u''-v'')] \geq 0 \qquad \text{q.o. in } \Omega$$

e quindi:

$$\int_\Omega -u''(v-u)\,dx + \int_\Omega Eu''(v''-u'')\,dx \geq \int_\Omega f[(v-u) - E(v''-u'')]\,dx$$

ossia:

$$\int_\Omega u'(v'-u')\,dx + \int_\Omega Eu''(v''-u'')\,dx \geq \int_\Omega f[(v-u) - E(v''-u'')]\,dx.$$

Ricordando le (19) e (20), la soprascritta relazione è appunto la (21).

Dim. (21) $\Rightarrow$ (4.1). Supponiamo che u sia soluzione della (21). Proveremo che u soddisfa alle (11), (12), (13) di pag. 269 e ciò equivale, per la (4.2), al fatto che u è soluzione del problema (4.1).

Essendo u elemento di K, u deve verificare la (11).

Allo scopo di provare la (12), sia φ un arbitrario elemento di $C_0^\infty(\Omega)$ tale che $\varphi \geq 0$. Applicando la (2.1) di pag. 260 con $c = \frac{1}{E}$, $g = \frac{\varphi}{E}$, si deduce che esiste $v_0 \in H$ tale da risultare:

(22) $\qquad v_0 - Ev_0'' = \varphi \qquad \text{q.o. in } \Omega$

e quindi, posto:

$$u + v_0 = v,$$

tenendo presente che $u \in K$ si ha:

$$v - Ev'' = u - Eu'' + \varphi \geq u - Eu'' \geq \psi + Ef \qquad \text{q.o. in } \Omega.$$

Pertanto $v \in K$. Ponendo allora nella (21) $v = u + v_0$ si ottiene:

$$a(u, v_0) \geq \langle F, v_0 \rangle ,$$

ossia:

$$\int_\Omega (u' v_0' + E u'' v_0'') \, dx \geq \int_\Omega f (v_0 - E v_0'') \, dx . \tag{23}$$

Essendo:

$$\int_\Omega u' v_0' \, dx = - \int_\Omega u'' v_0 \, dx ,$$

la (23) si scrive:

$$\int_\Omega -u'' (v_0 - E v_0'') \, dx \geq \int_\Omega f (v_0 - E v_0'') \, dx ,$$

ossia, per la (22):

$$\int_\Omega (-u'' - f) \varphi \geq 0 .$$

Poiché questa relazione sussiste $\forall \varphi \in C_0^\infty(\Omega)$ tale che $\varphi \geq 0$, ne segue che $-u'' - f \geq 0$ quasi ovunque in Ω [cfr. (12) pag. 150], ossia che vale la (12).

Resta da dimostrare che u soddisfa alla (13).

Applicando la (2.1) di pag. 260 con $c = \frac{1}{E}$ e $g = f + \frac{u}{E}$, si deduce che esiste un elemento v_1 di H tale da risultare:

$$v_1 - E v_1'' = \psi + E f \quad \text{q.o. in } \Omega . \tag{24}$$

Essendo v_1 un elemento di K in virtù della (24), si può porre nella (21) $v = v_1$. Si ottiene:

$$a(u, v_1 - u) \geq \langle F, v_1 - u \rangle ,$$

ossia:

$$\int_\Omega [u' (v_1' - u') + E u'' (v_1'' - u'')] \, dx \geq \int_\Omega f [(v_1 - u) - E (v_1'' - u'')] \, dx . \tag{25}$$

Essendo:

$$\int_\Omega u' (v_1' - u') \, dx = - \int_\Omega u'' (v_1 - u) \, dx ,$$

la (25) si scrive:

$$\int_\Omega -u'' [(v_1 - u) - E (v_1'' - u'')] \, dx \geq \int_\Omega f [(v_1 - u) - E (v_1'' - u'')] \, dx$$

ossia, per la (24):

$$\int_\Omega (-u'' - f)(-u + E u'' + \psi + E f) \, dx \geq 0$$

cioè:

$$\int_\Omega (u'' + f)(u - E u'' - \psi - E f) \, dx \geq 0 . \tag{26}$$

D'altra parte, per le (11) e (12) già dimostrate, il primo membro della (26) è ≤ 0, e quindi per la (26) stessa l'integrale a primo membro è nullo. Poiché la funzione integranda ha segno costante per le (11), (12), ne segue che essa è nulla q.o. in Ω.

La (13) è così dimostrata, e con essa l'assunto.

Dal teorema (4.3) consegue:

(4.4). Il problema (4.1) ammette, qualunque siano i dati f e ψ, una e una sola soluzione.

Dim. Per la (4.3) e per il teorema di Stampacchia [(1.1) pag. 255], basta constatare che la forma bilineare $a(u,v)$ definita dalla (19) di pag. 270 è continua e coercitiva, e che il funzionale (20) è continuo.

Ma ciò è stato già da noi dimostrato [cfr. es. 3 pag. 214 e es. 12 pag. 148].

La soluzione u del problema, essendo elemento dello spazio $H_0^1(\Omega) \cap H^2(\Omega)$, è una classe di funzioni definite quasi ovunque in Ω, ed è chiaro che una tale classe non è atta a rappresentare la configurazione di equilibrio del filo. Quest'ultima sarà ovviamente il diagramma del rappresentante $\tilde{u} \in C^{1,\frac{1}{2}}(\bar{\Omega})$ della classe u, esistente per il teorema di Sobolev [pag. 105].

E' interessante notare che, anche nel caso di un ostacolo "a punta", la configurazione di equilibrio è rappresentata dal diagramma di una funzione derivabile con derivata continua, mentre dell'analoga funzione relativa al problema del numero precedente è assicurata solo la continuità: in accordo con l'intuito fisico, il fatto che l'ostacolo sia deformabile impedisce che alle eventuali "punte" dell'ostacolo possano corrispondere delle discontinuità della derivata.

Osserviamo anche qui che, per la (1.2) di pag. 256, dal teor. (4.3) consegue un'altra interpretazione della soluzione u del problema. Infatti u è il punto che minimizza nel convesso K il funzionale:

$$\Phi(v) = \frac{1}{2}\int_\Omega (v'^2 + Ev''^2)\,dx - \int_\Omega f\cdot(v - Ev'')\,dx\,.$$

5. Un problema sulla trave incastrata.

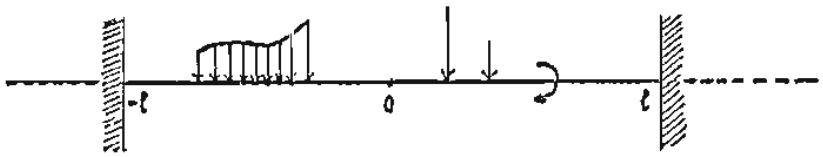
Poniamo il seguente problema.

Si consideri una trave di materiale omogeneo ed isotropo, a sezione variabile, incastrata rigidamente agli estremi e sottoposta ad un carico costituito da forze verticali, sia distribuite che concentrate, e da momenti concentrati. Nell'ipotesi che gli effetti dovuti al taglio siano trascurabili rispetto a quelli dovuti al momento flettente, determinare gli spostamenti verticali delle sezioni della trave.

Fissiamo l'asse x coincidente con l'asse della trave e l'origine nel punto medio di esa. Denotiamo con 2l la lunghezza della trave, con I(x) il momento d'inerzia, rispetto all'asse neutro, della sezione di ascissa x, e con B(x) la rigidità flessionale della trave:

$$B(x) = E\,I(x),$$

essendo E il modulo di Young del materiale della trave.



Per esaminare come conviene rappresentare le forze concentrate, nonché i momenti concentrati, osserviamo quanto segue.

Se la funzione f(x) rappresenta una forza distribuita sulla trave, e se $\varphi(x)$ rappresenta lo spostamento verticale della sezione della trave nel punto di ascissa x, il lavoro che la forza f deve compiere per produrre lo spostamento φ è dato da

$$L = \int_{-l}^{l} f(x)\varphi(x)\,dx, \tag{1}$$

e quindi, se $f \in L^1_{loc}(\Omega)$ con $\Omega =]-l,l[$ il funzionale:

$$L : \varphi \in C_0^\infty(\Omega) \to \int_{-l}^{l} f(x)\varphi(x)\,dx = \int_{-\infty}^{+\infty} f(x)\varphi(x)\,dx \tag{2}$$

può chiamarsi <u>il lavoro della forza f</u>.

Chiaramente sarebbe utile esprimere il lavoro mediante un funzionale dello stesso tipo anche nel caso che vi sia una sola forza F concentrata in un punto,

diciamo nel punto di ascissa x_0. Ma per rappresentare tale forza mediante una funzione $f(x)$ si dovrebbe porre $f(x)=F_0$ per $x=x_0$, $f(x)=0$ per $x\neq x_0$; l'integrale a secondo membro della (1) sarebbe allora uguale a zero, mentre il lavoro è

$$F_0\,\varphi(x_0) = f(x_0)\varphi(x_0).$$

L'esigenza di definire una funzione $f(x)$ per la quale il secondo membro della (1) sia uguale ad $f(x_0)\varphi(x_0)$ conduce a considerare la distribuzione di Dirac. Infatti, poichè risulta [cfr. (25) pag. 154]:

$$\int_{-\infty}^{+\infty} \delta(x-x_0)\,\varphi(x)\,dx = \varphi(x_0) \qquad \forall \varphi \in C_0^\infty(\Omega),$$

il lavoro si può esprimere ancora mediante il funzionale (2), sostituendo $f(x)$ con $F_0\,\delta(x-x_0)$. In tal modo un sistema di n forze $F_1, F_2, \dots, F_n$ concentrate rispettivamente nei punti di ascisse $x_1, \dots, x_n$ viene ad essere rappresentato da:

$$\sum_{i=1}^{n} F_i\,\delta(x-x_i) \quad . \tag{3}$$

E' opportuno però perfezionare tale conclusione. Ricordando che la distribuzione di Dirac relativa al punto x_i è la derivata nel senso delle distribuzioni della funzione scalino [cfr. (40) pag. 157]:

$$U_i : x \in R \longrightarrow \begin{cases} 0 & \text{se } x < x_i \\ 1 & \text{se } x \geq x_i \end{cases},$$

cioè che risulta:

$$\int_{-\infty}^{+\infty} \delta(x-x_i)\,\varphi(x)\,dx = -\int_{-\infty}^{+\infty} U_i(x)\,\varphi'(x)\,dx \qquad \forall \varphi \in C_0^\infty(\Omega),$$

il lavoro compiuto dal predetto sistema di forze concentrate è rappresentato dal funzionale L definito ponendo:

$$\langle L, \varphi \rangle = -\int_{-\infty}^{+\infty} \Big(\sum_{i=1}^{n} F_i\,U_i(x) \Big)\,\varphi'(x)\,dx \qquad \forall \varphi \in C_0^\infty(\Omega)$$

che è la derivata nel senso delle distribuzioni della funzione costante a tratti:

$$U(x) = \sum_{i=1}^{n} F_i\,U_i(x).$$

Conviene allora far riferimento a tale funzione invece che alla distribuzione (3), e rappresentare un sistema di forze concentrate mediante l'effetto che esso produce (il lavoro), cioè mediante la derivata nel senso delle distribuzioni di una funzione costante a tratti, o più generalmente di un elemento di $L^1_{loc}(\Omega)$.

Analogamente, se si considera un momento concentrato nel punto di ascissa x_i, di valore M_i, il lavoro della coppia relativo allo spostamento φ è

dato da $M_i \varphi'(x_i)$, e siccome risulta:

$$\varphi'(x_i) = \langle \delta_{x_i}, \varphi' \rangle = -\langle \delta'_{x_i}, \varphi \rangle = -\langle U_i'', \varphi \rangle ,$$

cioè:

$$\varphi'(x_i) = -\int_{-\infty}^{+\infty} U_i(x) \varphi''(x) dx ,$$

il lavoro complessivo è rappresentato dal funzionale M definito ponendo:

$$\langle M, \varphi \rangle = -\int_{-\infty}^{+\infty} \Big(\sum_{i=1}^{n} M_i U_i(x) \Big) \varphi''(x) dx \qquad \forall \varphi \in C_0^\infty(\Omega) ,$$

che è la derivata seconda nel senso delle distribuzioni della funzione costante a tratti:

$$U(x) = -\sum_{i=1}^{n} M_i U_i(x) .$$

Dunque un sistema di momenti concentrati si può rappresentare mediante la derivata seconda nel senso delle distribuzioni di una funzione costante a tratti, o più generalmente di un elemento di $L^1_{loc}(\Omega)$.

Orbene, posto $\Omega =]-\ell, \ell[$, rappresentiamo le forze distribuite mediante una funzione q_0, che supponiamo appartenente ad $L^2(\Omega)$, le forze concentrate mediante la derivata (nel senso delle distribuzioni) di una funzione $-q_1 \in L^2(\Omega)$, e i momenti concentrati mediante la derivata seconda (nel senso delle distribuzioni) di una funzione $q_2 \in L^2(\Omega)$. Pertanto il carico globale cui è sottoposta la trave è rappresentato dalla distribuzione:

$$q = q_0 - q_1' + q_2'' .$$

Supponiamo poi che la funzione $B(x)$ appartenga ad $L^\infty(\Omega)$ e sia essenzialmente positiva in senso stretto, cioè tale da risultare:

$$B(x) \geq c_0 > 0 \quad \text{q.o. in } \Omega \qquad (c_0 = \text{costante}) .$$

Com'è noto dalla Scienza delle costruzioni, detto $u(x)$ lo spostamento verticale della sezione di ascissa x deve risultare:

$$(4) \qquad (B u'')'' = q$$

$$(5) \qquad u(-\ell) = u(\ell) = 0 \quad , \quad u'(-\ell) = u'(\ell) = 0 .$$

Cercheremo la soluzione nello spazio $H_0^2(\Omega)$. Siccome per il teorema di Sobolev [pag. 105] gli elementi di $H_0^2(\Omega)$ hanno un rappresentante di classe $C^{1,\frac{1}{2}}(\overline{\Omega})$ che si annulla negli estremi di Ω insieme alla sua derivata prima, le condizioni (5) sono verificate automaticamente da tale rappresentante.

Naturalmente nella (4) u'' e' la derivata debole di u, mentre la derivata seconda di Bu'' va intesa nel senso delle distribuzioni.

Formuliamo dunque il seguente

PROBLEMA (5.1).

Assegnati q_0, q_1, q_2 elementi di $L^2(\Omega)$, e $B \in L^\infty(\Omega)$ con $B(x) \geq c_0 > 0$ q.o. in Ω, ricercare $u \in H_0^2(\Omega)$ tale che si abbia:

(6) $\qquad (Bu'')'' = q$ nel senso delle distribuzioni su Ω,

dove $q = q_0 - q_1' + q_2''$.

La (6) significa che:

$$\int_\Omega Bu''\varphi''dx = \langle q, \varphi \rangle \qquad \forall \varphi \in C_0^\infty(\Omega)$$

e il secondo membro e' dato da:

$$\langle q, \varphi \rangle = \langle q_0, \varphi \rangle - \langle q_1', \varphi \rangle + \langle q_2'', \varphi \rangle =$$

$$= \int_\Omega q_0\varphi dx + \int_\Omega q_1\varphi' dx + \int_\Omega q_2\varphi'' dx = \int_\Omega (q_0\varphi + q_1\varphi' + q_2\varphi'')dx \qquad \forall \varphi \in C_0^\infty(\Omega).$$

Pertanto la (6) esplicitamente si scrive:

(7) $$\int_\Omega Bu''\varphi''dx = \int_\Omega (q_0\varphi + q_1\varphi' + q_2\varphi'')dx \qquad \forall \varphi \in C_0^\infty(\Omega).$$

Posto, per ogni $u, v \in H_0^2(\Omega)$:

$$a(u,v) = \int_\Omega Bu''v''dx \quad , \quad \langle f, v \rangle = \int_\Omega (q_0 v + q_1 v' + q_2 v'')dx ,$$

siccome $C_0^\infty(\Omega)$ e' denso in $H_0^2(\Omega)$ [pag.104] la (7) e' equivalente all'equazione variazionale:

(8) $\quad u \in H_0^2(\Omega): \quad a(u,v) = \langle f, v \rangle \quad \forall v \in H_0^2(\Omega)$.

Orbene, la forma bilineare $a(u,v)$ e' continua e coercitiva [es.2 pag. 213], e il funzionale f e' continuo [es.11 pag.148]: pertanto, per il teorema di Lax-Milgram [(1.4) pag. 257] la (8) ammette una ed una sola soluzione.

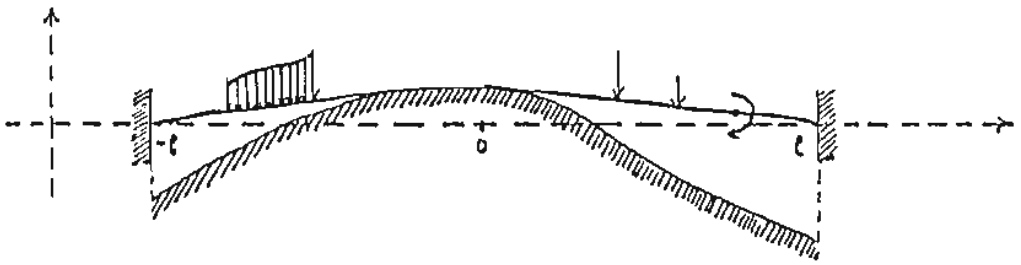
La soluzione e' un elemento dello spazio $H_0^2(\Omega)$: naturalmente, come per il problema del n.4, gli spostamenti verticali della

trave sono forniti dal rappresentante di classe $C^{1,\frac{1}{2}}(\bar{\Omega})$ della soluzione.

6. Un problema sulla trave incastrata e forzata su un ostacolo.

Consideriamo qui lo stesso problema del numero precedente, con l'ulteriore condizione che alla trave venga applicata una distorsione forzandola a poggiare su un ostacolo rigido. Il contatto trave-ostacolo si suppone privo di attrito.

Fissato l'asse x come nel n. 5 [pag. 274], introduciamo le stesse notazioni e conserviamo le stesse ipotesi. Quanto all'ostacolo, supponiamo che sia rappresentato dal diagramma di una funzione $\psi(x)$ continua nell'intervallo $[-\ell, \ell]$, tale che $\psi(-\ell) < 0$, $\psi(\ell) < 0$.



Anche qui cercheremo la soluzione nello spazio $H_0^2(\Omega)$, con $\Omega =]-\ell, \ell[$. Devono sussistere le stesse condizioni ai limiti [cfr. (5) pag. 276], che saranno automaticamente soddisfatte per l'appartenenza di u ad $H_0^2(\Omega)$, ed inoltre deve risultare $u \geq \psi$ in Ω, nel senso che tale disuguaglianza sia verificata dal rappresentante di u di classe $C^{1,\frac{1}{2}}(\bar{\Omega})$.

Quanto alla condizione $(Bu'')'' = q$, ora il carico globale deve essere sostituito con $q+\rho$, essendo ρ la reazione vincolare dell'ostacolo, la quale ha la direzione e il verso dell'asse y e quindi è positiva. Pertanto deve risultare $(Bu'')'' = q+\rho$, e quindi $(Bu'')'' \geq q$, naturalmente nel senso delle distribuzioni; in altri termini la distribuzione $(Bu'')'' - q$ deve essere ≥ 0, e ciò vuol dire che:

$$(1) \qquad \langle (Bu'')'' - q, \varphi \rangle \geq 0 \quad \forall \varphi \in C_0^\infty(\Omega), \ \varphi \geq 0.$$

Però nei punti in cui la trave non tocca l'ostacolo deve risultare $(Bu'')'' = q$, ossia, posto:

(2) $$\Omega' = \{x \in \Omega : u(x) > \psi(x) \text{ q.o.}\}$$

deve essere $(Bu'')'' = q$ nel senso delle distribuzioni su Ω'.

Formuliamo dunque il seguente

PROBLEMA (6.1).

Assegnati q_0, q_1, q_2 elementi di $L^2(\Omega)$, $B \in L^\infty(\Omega)$ con $B(x) \geq c_0 > 0$ q.o. in Ω, $\psi \in C^0(\bar{\Omega})$ con $\psi(-\ell) < 0$, $\psi(\ell) < 0$, ricercare $u \in H_0^2(\Omega)$ tale che si abbia:

(3) $$\begin{cases} (Bu'')'' \geq q & \text{nel senso delle distribuzioni su } \Omega \\ u \geq \psi & \text{q.o. in } \Omega \\ (Bu'')'' = q & \text{nel senso delle distribuzioni su } \Omega' \end{cases}$$

dove $q = q_0 - q_1' + q_2''$ e Ω' e' l'insieme (2).

Come nel n.5 poniamo, per $u, v \in H_0^2(\Omega)$:

$$a(u,v) = \int_\Omega Bu''v''\,dx \,, \qquad \langle f, v\rangle = \int_\Omega (q_0 v + q_1 v' + q_2 v'')\,dx \,.$$

Introduciamo quindi il sottoinsieme di $H_0^2(\Omega)$:

$$K = \{v \in H_0^2(\Omega) : v \geq \psi \text{ q.o.}\}$$

che e' non vuoto, chiuso e convesso. Ciò si dimostra come si e' fatto per l'insieme (4) di pag. 262 (*).

Orbene l'esistenza e l'unicità per il problema (6.1) seguiranno dal fatto che:

(6.2). Il problema (6.1) e' equivalente alla disequazione variazionale:

(4) $$u \in K : \quad a(u, v-u) \geq \langle f, v-u\rangle \quad \forall v \in K.$$

Il teorema si dimostra con gli stessi procedimenti seguiti per la (3.1) di pag. 263, che si riferisce al problema espresso dalle (3) di pag. 261, che e' dello stesso tipo del problema (6.1) [nel quale figura $(Bu'')''$ invece di $-u''$].

Pertanto riportiamo solo le tappe fondamentali della dimostrazione.

Dim. (3) $\Rightarrow$ (4). Per la seconda delle (3), $u \in K$. Per la prima delle (3), la distribuzione $(Bu'')'' - q$ e' positiva, e quindi esiste una misura μ, definita nell'insieme delle parti misurabili di Ω, che assume valore finito sui compatti, ed e' tale da risultare:

(*) Per la dimostrazione del fatto che $K \neq \emptyset$, qui occorre costruire un elemento di $H_0^2(\Omega)$ invece che di $H_0^1(\Omega)$: basterà sostituire $(\delta^2 - x^2)^2$ con $(\delta^2 - x^2)^3$ e $(\delta^2 - x^2)^2$ con $(\delta^2 - x^2)^3$.

$$(5) \qquad \langle (Bu'')'' - q, \varphi \rangle = \int_\Omega \varphi \, d\mu \quad \forall \varphi \in C_0^\infty(\Omega).$$

Fissato $v \in K$, e detta $\{\varphi_n\}$ una successione di elementi di $C_0^\infty(\Omega)$ convergente a $v-u$ in $H_0^1(\Omega)$, dalla (5) si deduce:

$$(6) \qquad a(u, \varphi_n) - \langle f, \varphi_n \rangle = \int_\Omega \varphi_n \, d\mu \quad \forall n \in N.$$

Detti $\tilde{u}$ e $\tilde{v}$ i rappresentanti continui di u e v, per il teorema di Sobolev la successione $\{\varphi_n\}$ converge a $\tilde{v}-\tilde{u}$ uniformemente; d'altra parte la terza delle (3) consente di stabilire che $\mu(\Omega')=0$, da cui si deduce $\tilde{v}-\tilde{u} \geq 0$ μ-quasi ovunque in Ω, e che $\mu(\Omega) < +\infty$.

Dopo ciò dalla (6) passando al limite per $n \to \infty$ si ottiene la (4), tenuto conto che a secondo membro è lecito il passaggio al limite sotto il segno di integrale e che l'integrale del limite $\tilde{v}-\tilde{u}$ è ≥ 0.

Dim. (4) $\Rightarrow$ (3). La seconda delle (3) segue dal fatto che $u \in K$. La prima si ottiene dalla (4) ponendo $v = u + \varphi$ con $\varphi \in C_0^\infty(\Omega)$, $\varphi \geq 0$.

Per stabilire la terza delle (3), fissato $\varphi \in C_0^\infty(\Omega')$ si costruisce un numero $\varepsilon > 0$ in modo che $u - \varepsilon\varphi$ appartenga al convesso K. Ponendo poi in (4) $v = u - \varepsilon\varphi$ si perviene facilmente all'assunto.

Dalla (6.2) consegue che il problema (6.1) ammette una ed una sola soluzione: basta applicare il teorema di Stampacchia [(1.1) pag. 255], tenuto conto di quanto richiamato a pag. 277 dopo la (8).

La soluzione del problema fisico considerato è fornita dal rappresentante di classe $C^{1,\frac{1}{2}}(\bar{\Omega})$ della soluzione del problema (6.1): l'esistenza di tale rappresentante è assicurata dal teorema di Sobolev [pag. 105].

In virtù del teor. (1.2) di pag. 256, abbiamo anche stabilito che nel convesso K il funzionale:

$$(7) \qquad \Phi(v) = \frac{1}{2} \int_\Omega B v''^2 \, dx - \int_\Omega (q_0 v + q_1 v' + q_2 v'') \, dx,$$

definito in $H_0^2(\Omega)$, è dotato di minimo, e che il valore minimo è assunto in un unico punto di K.

Ricordiamo [teor. (1.3) pag. 257] che un elemento u del convesso K è soluzione della disequazione variazionale (4) se e solo se risulta:

$$(8) \qquad d\Phi(u)(v-u) \geq 0 \quad \forall v \in K.$$

Di conseguenza per la soluzione u del problema (6.1), cioè per il punto che minimizza il funzionale (10) nel convesso K, si ha:

$$d\Phi(u) \geq 0 \quad \text{in } H_0^2(\Omega), \tag{9}$$

ossia:

$$d\Phi(u)z \geq 0 \quad \forall z \in H_0^2(\Omega), z \geq 0 \ (^1). \tag{9'}$$

Notiamo che il funzionale (7) e il differenziale $d\Phi(v)$ hanno significato anche per $v \in H^2(\Omega)$, e quindi se si suppone $\psi \in H^2(\Omega)$ (^{2}) ha senso considerare il differenziale $d\Phi(\psi)$. Orbene, se tale differenziale e' ≥ 0 in $H_0^2(\Omega)$, cioe', come facilmente si verifica (^{3}), se risulta:

$$(B\psi')' \geq q \quad \text{nel senso delle distribuzioni su } \Omega, \tag{10}$$

si puo' dire che la superficie di contatto trave-ostacolo e' un insieme connesso. E' questa un'interpretazione del seguente teorema:

(6.3). Detta u la soluzione del problema (6.2), sia Ω_0 l'insieme costituito dalle ascisse dei punti di contatto trave-ostacolo, cioe' l'insieme degli $x \in \Omega$ tali che $u(x) = \psi(x)$.

Se ψ appartiene ad $H^2(\Omega)$ e verifica la (10), allora Ω_0 e' un intervallo.

Dim. Si tratta di far vedere che qualunque siano $a, b \in \Omega_0$, con $a < b$, si ha $[a,b] \subseteq \Omega_0$. Dimostreremo che la funzione $w(x)$ definita in $\bar{\Omega}$ ponendo:

$$w(x) = \begin{cases} \psi(x) & \text{se } x \in [a,b] \\ u(x) & \text{se } x \notin [a,b] \end{cases}$$

coincide con u (4) : ne seguira' che $u(x) = \psi(x)$ $\forall x \in [a,b]$, ossia che $[a,b] \subset \Omega_0$.

In primo luogo occorre far vedere che w e' un elemento di $H_0^2(\Omega)$.

Per l'ipotesi che a e b appartengono ad Ω_0 si ha:

$$u(a) = \psi(a) \quad , \quad u(b) = \psi(b), \tag{11}$$

e perciò la funzione w e' continua in $\bar{\Omega}$.

D'altra parte, i punti a e b sono necessariamente interni all'intervallo $[-l, l]$, come tutti i punti di Ω_0: infatti, essendo $u(-l) = u(l) = 0$ e $\psi(-l) < 0$, $\psi(l) < 0$,

---

(^{1}) Infatti qualunque sia $z \in H_0^2(\Omega)$, $z \geq 0$, il punto $u+z$ appartiene al convesso K (in quanto $u+z \geq u \geq \psi$) : ponendo allora nella (8) $v = u+z$, si ottiene $d\phi(u)z \geq 0$.

(^{2}) Precisamente, siccome $\psi \in C^0(\bar{\Omega})$: se si suppone che ψ sia il rappresentante continuo di un elemento di $H^2(\Omega)$, denotato ancora con ψ.

(^{3}) A tale scopo si tenga presente che, se $z \in H_0^m(\Omega)$, $z \geq 0$, esiste una successione $(\varphi_n)_{n \in \mathbb{N}}$ di elementi di $C_0^\infty(\Omega)$ convergente a z in $H_0^m(\Omega)$ tale che $\varphi_n \geq 0$ $\forall n \in \mathbb{N}$. Il risultato sussiste in opportune ipotesi su $\Omega \subset \mathbb{R}^n$, che sono certamente verificate nel nostro caso, essendo Ω un intervallo limitato di $\mathbb{R}$.

(^{4}) Qui e nel seguito facciamo riferimento ai rappresentanti di u e ψ di classe $C^{1,\frac{1}{2}}(\bar{\Omega})$, che denotiamo ancora con u e ψ.

gli estremi $-\ell, \ell$ dell'intervallo non appartengono ad Ω_0. Allora dal fatto che $u \in K$, cioè dalla disuguaglianza:

$$u(x) \geq \psi(x) \quad \forall x \in \Omega \tag{12}$$

consegue:

$$u'(a) = \psi'(a) \quad , \quad u'(b) = \psi'(b) \, , \tag{13}$$

perchè se fosse, ad esempio, $u'(a) - \psi'(a) > 0$, la funzione $u(x) - \psi(x)$ sarebbe in a strettamente crescente e quindi, essendo nulla in a con a interno all'intervallo, assumerebbe valori negativi in un intorno sinistro di a, il che contraddice la (12).

In virtù delle (13), anche la funzione w' è continua in $\bar{\Omega}$.

Allo scopo di far vedere che esiste la derivata seconda debole di w, poniamo per $x \in \bar{\Omega}$:

$$w^*(x) = \begin{cases} \psi''(x) & \text{se } x \in [a,b] \\ u''(x) & \text{se } x \notin [a,b] \end{cases}$$

e dimostriamo che w^* è un rappresentante della derivata seconda debole di w, cioè che risulta:

$$\int_{-\ell}^{\ell} w^* \varphi \, dx = \int_{-\ell}^{\ell} w \varphi'' dx \qquad \forall \varphi \in C_0^\infty(\Omega) \, . \tag{14}$$

Fissato un elemento φ di $C_0^\infty(\Omega)$, denotiamo con c un numero positivo tale che l'intervallo $[-c,c]$ sia contenuto in Ω, e contenga nel suo interno il supporto di φ nonchè $[a,b]$. Si ha:

$$\int_{-\ell}^{a} w^* \varphi \, dx = \int_{-c}^{a} w^* \varphi \, dx = \int_{-c}^{a} u'' \varphi \, dx = [u' \varphi]_{-c}^{a} - \int_{-c}^{a} u' \varphi' dx \quad (*) =$$

$$= u'(a) \varphi(a) - \int_{-c}^{a} u' \varphi' dx = u'(a) \varphi(a) - [u \varphi']_{-c}^{a} + \int_{-c}^{a} u \varphi'' dx = u'(a) \varphi(a) - u(a) \varphi'(a) + \int_{-\ell}^{a} u \varphi'' dx \, ,$$

e quindi:

$$\int_{-\ell}^{a} w^* \varphi \, dx = u'(a) \varphi(a) - u(a) \varphi'(a) + \int_{-\ell}^{a} w \varphi'' dx \, . \tag{15}$$

Analogamente si trova:

$$\int_{b}^{\ell} w^* \varphi \, dx = -u'(b) \varphi(b) + u(b) \varphi'(b) + \int_{b}^{\ell} w \varphi'' dx \quad , \tag{16}$$

e d'altra parte si ha, tenuto conto che in $[a,b]$ risulta $\psi = w$:

$$\int_{a}^{b} w^* \varphi \, dx = \int_{a}^{b} \psi'' \varphi \, dx = [\psi' \varphi]_{a}^{b} - [\varphi' \psi]_{a}^{b} + \int_{a}^{b} w \varphi'' dx \, . \tag{17}$$

Dalle (15), (16), (17), sommando membro a membro e tenendo conto delle (11) e (13), consegue

---

(*) Per ogni intervallo $[\alpha,\beta] \subset \Omega$ è applicabile la regola di integrazione per parti:

$$\int_{\alpha}^{\beta} u'' \varphi \, dx = [u' \varphi]_{\alpha}^{\beta} - \int_{\alpha}^{\beta} u' \varphi' dx \, . \tag{$\varkappa$}$$

Infatti un elemento z di $H^1(\Omega)$ essendo dotato di derivata debole ha un rappresentante assolutamente continuo in ogni intervallo compatto contenuto in Ω [cfr. pag. 60], sicchè qualunque sia $[\alpha,\beta] \subset \Omega$ si ha [cfr. pag. 47]:

$$\int_{\alpha}^{\beta} z'(x) dx = z(\beta) - z(\alpha) \, .$$

Da tale relazione ponendo $z = u' \varphi$ si deduce la ($\varkappa$).

la (14). Dunque w^* è un rappresentante di w'', e siccome w^* è di quadrato sommabile in $[a,b]$ e in $\Omega-[a,b]$ (giacchè tali sono ψ'' e u''), quindi in Ω, ne segue che $w \in H^2(\Omega)$.

Poichè w e w' si annullano negli estremi di Ω [in quanto $w(x)=u(x)\ \forall x \notin [a,b]$], per il teorema di Sobolev [pag.105] si conclude che $w \in H_0^2(\Omega)$.

E' evidente che $w \in K$ [o $w(x) \circ \psi(x)$, oppure $w(x) \equiv u(x) \geq \psi(x)$]. Ponendo allora nella disequazione variazionale (4) $v=w$ si ottiene:

$$a(u, w-u) \geq \langle f, w-u \rangle \ . \tag{18}$$

D'altra parte per l'ipotesi (10), tenuto conto che $u-w \in H_0^2(\Omega)$ e $u-w \geq 0$ (o risulta $u(x)-w(x)=0$ oppure $w(x)=\psi(x) \leq u(x)$), si ha [1]:

$$a(\psi, u-w) \geq \langle f, u-w \rangle$$

cioè, essendo $a(\psi, u-w) = a(w, u-w)$ [2]:

$$a(w, u-w) \geq \langle f, u-w \rangle ,$$

ossia:

$$a(-w, w-u) \geq \langle f, u-w \rangle \ . \tag{19}$$

Da (18) e (19), sommando membro a membro, consegue:

$$a(u-w, w-u) \geq 0$$

ossia:

$$a(u-w, u-w) \leq 0 \ . \tag{20}$$

Siccome la forma bilineare $a(u,v)$ è coercitiva [cfr.2 pag.213], la (20) implica $w=u$ [cfr. (5) pag.255] e con ciò l'assurdo è dimostrato.

— — — — — — — — — —

---

[1] La (10) comporta che $\forall v \in H_0^2(\Omega)$, $v \geq 0$, risulta $a(\psi, v) - \langle f, v \rangle \geq 0$.

[2] Essendo:

$$a(\psi, u-w) = \int_\Omega B\psi''(u''-w'')\,dx \quad , \quad a(w, u-w) = \int_\Omega Bw''(u''-w'')\,dx$$

basta ricordare che in $[a,b]$ si ha $\psi''(x) \equiv w^*(x)$ e in $\Omega-[a,b]$ risulta $u''(x)-w^*(x)=0$ (w^* è il rappresentante continuo di w'').

Capitolo X - Cenni di Analisi non lineare.

1. Punti fissi

Assegnata una funzione F, definita in un insieme S e a valori in S, consideriamo l'equazione nell'incognita $x \in S$:

$$F(x) = x \,. \tag{1}$$

Se esistono soluzioni della (1), ciascuna di esse si chiama un *elemento unito di F*, oppure un *punto fisso di F*; conseguentemente, per esprimere il fatto che la (1) ammette soluzioni, si dice che F è *dotata di elementi uniti*, oppure che F *ha almeno un punto fisso*.

Numerosi problemi di Analisi possono essere ricondotti allo studio di una equazione del tipo (1): perciò la teoria dei punti fissi ha avuto un ampio sviluppo. In questo numero ci proponiamo di dare un cenno di alcuni risultati principali, che sono suscettibili di svariate spplicazioni.

La proposizione che segue, del tutto ovvia, esprime il cosiddetto *metodo delle approssimazioni successive*:

(1.1) *Siano: S uno spazio topologico di Hausdorff, F un'applicazione continua di S in S.*

Scelto un punto $x_0 \in S$, *si ponga:*

$$x_n = F(x_{n-1}) \quad \forall n \in N \,. \tag{2}$$

Se esiste il limite:

$$\lim x_n \tag{3}$$

diciamolo $\bar{x}$, *allora* $\bar{x}$ *è un punto fisso di F.*

- - - - - - - - - - - - - -

Dim. Infatti dalla (2), passando al limite per $n \to \infty$ e tenendo conto della continuità di F, consegue [cfr. (1.10) pag. 110] :

$$\bar{x} = F(\bar{x}) \,.$$

- - - - - - - - - - - - - - - - - - -

Allo scopo di formulare una importante condizione che assicura l'esistenza del limite (3), introduciamo la seguente definizione.

Sia (S,d) uno spazio metrico. Si dice che l'applicazione:

$$F : S \to S$$

è una contrazione di (S,d), quando esiste un numero positivo K, minore di 1, tale che:

$$(4) \qquad d(F(x),F(y)) \leq K\, d(x,y) \qquad \forall x,y \in S .$$

Sussiste il seguente teorema di Banach-Caccioppoli (principio delle contrazioni):

(1.2) Ogni contrazione F di uno spazio metrico (S,d) ammette al più un punto fisso.

Se lo spazio (S,d) è completo, allora F ha un punto fisso (e uno solo). Scelto comunque $x_0 \in S$, e posto:

$$(5) \qquad x_n = F(x_{n-1}) \qquad \forall n \in N ,$$

il punto fisso $\bar{x}$ di F è il limite della successione $(x_n)_{n\in N}$, e risulta:

$$(6) \qquad d(x_n,\bar{x}) \leq \frac{d(x_0,F(x_0))}{1-K}\, K^n \qquad \forall n \in N ,$$

dove K è il numero che figura in (4).

Si osservi che la (6) fornisce, per ogni $n \in N$, una valutazione dell'errore che si commette se si assume x_n come punto fisso di F.

- - - - - - - - - - - - - -

Dim. Se $\bar{x}$ e $\bar{x}'$ sono punti fissi di F, dalla (4) consegue:

$$d(\bar{x},\bar{x}') \leq K\, d(\bar{x},\bar{x}') .$$

Se fosse $\bar{x} \neq \bar{x}'$, da tale relazione si dedurrebbe che $K \geq 1$, contro l'ipotesi.

Allo scopo di dimostrare la seconda parte del teorema, proviamo dapprima che per ogni $n \in N$ risulta:

$$(*) \qquad d(x_n, x_{n+1}) \leq K^n\, d(x_0, F(x_0)) .$$

La (*) è vera per n=1, in quanto:

$$d(x_1,x_2) = d(F(x_0),F(x_1)) \leq K\, d(x_0,x_1) = K\, d(x_0,F(x_0)) ;$$

d'altro canto, se la (*) è vera per n=m, essa è vera altresì per $n = m+1$, in quanto si ha:

$$d(x_{m+1},x_{m+2}) = d(F(x_m),F(x_{m+1})) \leq K\, d(x_m,x_{m+1}) \leq K^{m+1}\, d(x_0,F(x_0)) .$$

Pertanto la (k) è vera per ogni $n \in N$. Ne segue che, qualunque siano $n \in N$ e $p \in N$, risulta:

$$d(x_n, x_{n+p}) \leq \sum_{i=0}^{p-1} d(x_{n+i}, x_{n+i+1}) \leq \sum_{i=0}^{p-1} k^{n+i} d(x_0, F(x_0)) =$$

$$= k^n (1 + k + \cdots + k^{p-1}) d(x_0, F(x_0))$$

e quindi:

$$(**) \qquad d(x_n, x_{n+p}) < \frac{k^n}{1-k} d(x_0, F(x_0)) \quad \forall p, n \in N.$$

Poichè la successione $(k^n)_{n \in N}$ tende a 0 in quanto $0 < k < 1$, dalla (**) si deduce che $(x_n)_{n \in N}$ è una successione di Cauchy: pertanto, essendo (S,d) completo, tale successione è convergente [e tende al punto fisso $\bar{x}$ di F, in virtù di (1.1)] .

Inoltre dalla (xx) , passando al limite per $p \to \infty$, si deduce la (C).

- - - - - - - - - - - - - - - -

Sia (S,d) uno spazio metrico, ed F un'applicazione di S in S.

Si dice che F è <u>non espansiva</u>, quando risulta:

$$d(F(x), F(y)) \leq d(x,y) \quad \forall x, y \in S.$$

Le contrazioni di (S,d) , nonchè le isometrie di S in S, sono particolari applicazioni non espansive; evidentemente, se S è uno spazio di Hilbert e H è un sottospazio vettoriale di S chiuso in S, il proiettore P_H di S su H è un'applicazione non espansiva (1).

Anche se (S,d) è completo, le applicazioni non espansive di S in S non hanno necessariamente punti fissi, a differenza di quanto accade per le contrazioni (2).

In proposito sussiste il seguente <u>teorema di Browder</u>, che ci limitiamo ad enunciare:

---

(1) Infatti risulta [cfr.(18) pag.120] :

$$\| P_H x - P_H y \| = \| P_H (x-y) \| \leq \| x - y \| \quad \forall x, y \in S.$$

(2) Ad esempio, se S è uno spazio di Banach e $x_0 \neq 0$ un punto di S, l'applicazione:

$$F : x \in S \to x + x_0$$

è non espansiva, e non ha alcun punto fisso.

(1.3) Siano: S uno spazio di Hilbert, X un sottoinsieme non vuoto di S. Se X è chiuso, convesso e limitato, allora ogni applicazione di X in X, non espansiva, ha almeno un punto fisso.

Il principio delle contrazioni, nonchè il teorema di Browder ora enunciato, pur essendo utilissimi in numerose questioni, si riferiscono a funzioni di tipo molto particolare, quali sono le contrazioni e le applicazioni non espansive. Di notevole importanza è perciò il seguente teorema di Schauder, di cui non riportiamo la dimostrazione:

(1.4) Siano: S uno spazio vettoriale normato, X un sottoinsieme non vuoto di S, ed F un'applicazione continua di X in X.
Se X è convesso, e se $F(X)$ è precompatto, esiste almeno un punto fisso di F.

Il teorema fu dimostrato da Schauder nell'ulteriore ipotesi che S sia completo, e che X sia chiuso. La dimostrazione di (1.4) si trova nel volume "Fixed point theorems" di D.M. Smart (Cambridge Univ. Press).

Il teorema di Schauder è una generalizzazione del classico teorema di Brouwer, il quale asserisce che, detto Γ il cerchio chiuso di R^n col centro nell'origine e raggio 1, ogni applicazione continua di Γ in Γ ha almeno un punto fisso.

Un'estensione del teorema di Schauder agli spazi vettoriali topologici è il seguente teorema di Tychonoff:

(1.5) Siano: S uno spazio vettoriale topologico localmente convesso, X un sottoinsieme non vuoto di S, ed F un'applicazione continua di X in X.
Se X è convesso e compatto, esiste almeno un punto fisso di F.

2. Differenziale.

Consideriamo una funzione:

$$F : \Omega \subseteq S \to S_1 ,$$

con S, S_1 spazi vettoriali topologici ed Ω aperto.

Si dice che F è <u>differenziabile secondo Gateaux</u>, o <u>in senso debole</u>, <u>nel punto</u> $x_0 \in \Omega$, quando per ogni $y \in S$ esiste il limite:

$$(1) \qquad \lim_{t \to 0} \frac{F(x_0+ty)-F(x_0)}{t} \qquad (1)$$

ed inoltre tale limite, come funzione di y, è un operatore lineare e continuo (di S in S_1). In tale ipotesi, l'operatore lineare e continuo di S in S_1 che ad ogni $y \in S$ associa il limite (1) si denota con $dF(x_0)$, e si chiama <u>il differenziale</u> della funzione F <u>nel punto</u> x_0.

In particolare se $S_1 = R$, cioè se F è un funzionale reale, il differenziale $dF(x_0)$ è un elemento dello spazio S^* duale topologico di S.

Nel caso che F sia una funzione reale di n variabili reali, la differenziabilità in senso debole comporta che F è derivabile in x_0 secondo qualunque direzione (quindi dotata in x_0 di tutte le derivate parziali prime). Infatti, detto r_y l'asse di R^n di versore y, per ogni punto x del tipo $x = x_0 + ty$ il numero t rappresenta la misura relativa del segmento orientato di estremi x ed x_0, diciamola $\overline{x x_0}$, e quindi si ha:

$$\lim_{t \to 0} \frac{F(x_0+ty)-F(x_0)}{t} = \lim_{\substack{x \to x_0 \\ x \in r_y}} \frac{F(x)-F(x_0)}{\overline{x x_0}} = \frac{\partial F}{\partial r_y}(x_0) \,,$$

sicché per ogni $y \in R^n$ di norma unitaria risulta:

$$< dF(x_0), y > = \frac{\partial F}{\partial r_y}(x_0)$$

e in particolare, detto e_i il versore dell'asse x_i:

$$< dF(x_0), e_i > = \frac{\partial F}{\partial x_i}(x_0).$$

Di conseguenza, siccome per ipotesi il limite (1) come funzione di y è un funzionale lineare (e continuo), ossia del tipo:

$$< dF(x_0), y > = c_1 y_1 + \cdots + c_n y_n \qquad \forall y = (y_1, \ldots, y_n) \in R^n,$$

si ha necessariamente $c_i = \frac{\partial F}{\partial x_i}(x_0)$; quindi $dF(x_0)$ è il differenziale in senso classico:

$$dF(x_0) : y \in R^n \longrightarrow \frac{\partial F}{\partial x_1}(x_0) y_1 + \cdots + \frac{\partial F}{\partial x_n}(x_0) y_n \,,$$

---

(1) Osserviamo che per ogni fissato y, se si considera la funzione di una variabile reale a valori in S_1:

$$\Phi(t) = F(x_0+ty) \,,$$

il limite (1) rappresenta $\Phi'(0)$, cioè la derivata di $\Phi(t)$ nel punto $t=0$.

e risulta:

$$(2) \qquad \langle dF(x_0), y \rangle = (\operatorname{grad} F(x_0), y) \qquad \forall y \in R^n.$$

Se S è uno spazio di Hilbert ed $S_1 = R$, per il teorema di Riesz [cfr. (7.1) pag. 177] esiste un unico $w \in S$ tale da risultare:

$$\langle dF(x_0), y \rangle = (w, y) \qquad \forall y \in S .$$

Poichè nel caso $S = R^n$ l'elemento w è il gradiente di F nel punto x_0 [cfr. (2)], anche nel caso generale w si chiama il gradiente di F nel punto x_0, e si denota con $\nabla F(x_0)$ o con $F'(x_0)$. Pertanto $F'(x_0)$ è l'elemento di S per il quale risulta:

$$\langle dF(x_0), y \rangle = (F'(x_0), y) \qquad \forall y \in S .$$

Se si identifica lo spazio di Hilbert S con il suo duale, il differenziale $dF(x_0)$ viene ad identificarsi col gradiente $F'(x_0)$, elemento di S. Da ciò deriva la consuetudine, qualunque siano gli spazi vettoriali topologici S, S_1, di denotare anche con $F'(x_0)$ il differenziale di una funzione $F : \Omega \subseteq S \to S_1$ differenziabile secondo Gateaux nel punto $x_0 \in \Omega$.

D'ora in avanti supporremo che S ed S_1 siano spazi vettoriali normati, sicchè $dF(x_0)$ è un elemento di $\mathcal{L}(S, S_1)$.

Nel caso che F è una funzione reale di n variabili reali, differenziabile in x_0 in senso debole, non è detto che la differenza tra l'incremento $F(x_0+y) - F(x_0)$ e il differenziale $dF(x_0)$ calcolato in y, sia un infinitesimo di ordine superiore rispetto a $\|y\|_{R^n}$, cioè non è detto che F sia continua in x_0 ed ivi differenziabile in senso classico. Anche nel caso generale, se F è differenziabile in x_0 secondo Gateaux, per la differenza $r(y)$ tra l'incremento e il differenziale:

$$r(y) = [\, F(x_0+y) - F(x_0) \,] - \langle dF(x_0), y \rangle \qquad (1)$$

non è detto che si abbia:

$$(3) \qquad \lim_{y \to 0} \frac{r(y)}{\|y\|} = 0 .$$

Orbene, se è verificata la (3) si dice che nel punto x_0 la funzione

(1) Naturalmente definita nell'insieme degli $y \in S$ tali che $x_0 + y \in \Omega$

F e' differenziabile in senso forte, o secondo Frèchet.

Abbiamo introdotto la differenziabilita' in x_0 secondo Frechet come la proprieta' che F sia differenziabile in x_0 in senso debole ed inoltre soddisfi alla condizione (3). Peraltro la definizione si puó porre indipendentemente dalla nozione di differenziabilita' in senso debole, dicendo che F e' differenziabile in x_0 secondo Fréchet quando esiste un operatore lineare e continuo $A : S \to S_1$ tale che si abbia:

$$\lim_{y \to 0} \frac{\| F(x_0+y) - F(x_0) - Ay \|}{\|y\|} = 0 . \tag{4}$$

E' facile allora verificare che di conseguenza F e' differenziabile in x_0 secondo Gateaux, e che l'operatore A e' necessariamente il differenziale $dF(x_0)$.

Infatti, posto $r(y) = F(x_0+y) - F(x_0) - Ay$, si ha:

$$\frac{F(x_0+ty)-F(x_0)}{t} = \frac{A(ty)+r(ty)}{t} = Ay + \frac{r(ty)}{t}$$

e quindi, essendo:

$$\lim_{t\to 0} \left\| \frac{r(ty)}{t} \right\| = \|y\| \lim_{t\to 0} \frac{\|r(ty)\|}{\|ty\|} = \|y\| \lim_{x\to 0} \frac{\|r(x)\|}{\|x\|} = 0$$

risulta:

$$\lim_{t\to 0} \frac{F(x_0+ty)-F(x_0)}{t} = Ay \qquad \forall y \in S .$$

ossia:

$$Ay = \langle dF(x_0), y \rangle \quad \forall y \in S .$$

Evidente e' il teorema:

(2.1). Siano: S, S_1 spazi vettoriali normati, Ω un aperto di S, F e G applicazioni di S in S_1. Per ogni $x_0 \in \Omega$ si ha che:

Se F e G sono differenziabili nel punto x_0 in senso debole [forte], anche $F+G$ e' differenziabile in x_0 in senso debole [forte], e risulta:

$$d(F+G)(x_0) = dF(x_0) + dG(x_0) .$$

Se F e' differenziabile in x_0 in senso debole [forte], tale e' anche λF qualunque sia $\lambda \in R$, e risulta:

$$d(\lambda F)(x_0) = \lambda\, dF(x_0) .$$

Il teorema seguente si dimostra con lo stesso procedimento dell'Analisi classica:

(2.2). Siano: S, S_1, S_2 spazi vettoriali normati, Ω un aperto di S, Ω_1 un aperto di S_1. Date le applicazioni continue:

$$F: \Omega \to \Omega_1 \quad , \quad G: \Omega_1 \to S_2 \ ,$$

si consideri un punto $x_0 \in \Omega$ e si ponga $y_0 = F(x_0) \in \Omega_1$.

Se F è differenziabile in x_0 in senso debole [forte] e G è differenziabile in y_0 in senso forte, la funzione composta $G \circ F$ è differenziabile in x_0 in senso debole [forte], e risulta:

$$d(G \circ F)(x_0) = dG(y_0) \cdot dF(x_0).$$

Osserviamo che:

(2.3) Qualunque siano gli spazi vettoriali normati S, S_1 si ha quanto segue.

I. Per ogni $y_0 \in S_1$ l'applicazione costante:

$$F: x \in S \to y_0$$

è differenziabile in senso forte in ogni punto $x_0 \in S$, e risulta:

$$dF(x_0) = 0 \ .$$

II. Ogni operatore lineare e continuo $L: S \to S_1$ è differenziabile in senso forte in ogni punto $x_0 \in S$, e risulta:

$$dF(x_0) = L \ .$$

III. Se S è uno spazio pre-hilbertiano reale, il funzionale:

$$F: x \in S \to \|x\|^2$$

è differenziabile in senso forte in ogni punto $x_0 \in S$, e risulta:

$$dF(x_0) = (2x_0, \cdot) \ .$$

IV. Se S è uno spazio pre-hilbertiano reale, e se $a(x,y)$ è

una forma bilineare continua in $S \times S$, il funzionale F definito ponendo:

$$F(x) = a(x,x)$$

è differenziabile in senso forte in ogni punto $x_0 \in S$, e risulta:

$$dF(x_0) = a(x_0,\cdot) + a(\cdot,x_0)$$

— — — — — — — — —

Dim. I. Basta osservare che la (4) di pag. 290 è banalmente verificata assumendo A identicamente nullo.

Dim II. Per la funzione $F = L$ la citata (4) è banalmente verificata se si assume $A = L$.

Dim. III. Si tratta di un caso particolare della IV.

Dim. IV. Essendo:

$$F(x_0+y) - F(x_0) = a(x_0+y, x_0+y) - a(x_0,x_0) =$$

$$= a(x_0,y) + a(y,x_0) + a(y,y)\,,$$

la (4) di pag. 290 è verificata con

$$A = a(x_0,\cdot) + a(\cdot,x_0)\,.$$

Infatti, essendo $|a(y,y)| \le l\,\|y\|^2$ con l costante (per la supposta continuità di $a(x,y)$), si ha:

$$\|F(x_0+y) - F(x_0) - Ay\| = \|a(y,y)\| \le l\,\|y\|^2$$

donde la (4).

— — — — — — — — — —

Con riferimento alla prop. IV, se la forma bilineare $a(x,y)$ è anche simmetrica si ha:

$$dF(x_0) = 2\,a(x_0,\cdot)$$

e si ottiene una precisazione del risultato enunciato nella proposiz. 2) di pag. 257.1, in quanto la differenziabilità è ivi intesa in senso debole.

Una generalizzazione del concetto di differenziale sarà accennata nell'Appendice, a pag. 326.

Come nell'Analisi classica, una condizione sufficiente affinche' una funzione F sia differenziabile in senso forte in un punto $x_0 \in \Omega$ e' che F sia differenziabile in senso debole in un intorno I di x_0 contenuto in Ω, e che inoltre la funzione $x \in I \to dF(x)$ sia continua in x_0. Per dimostrare tale asserzione occorre premettere il seguente teorema.che costituisce una estensione del teorema di Lagrange:

(2.4) Si supponga che la funzione

$$F : \Omega \subseteq S \to S_1$$

con S, S_1 *spazi vettoriali normati, sia differenziabile in senso debole in ogni punto del segmento* $\sum$ *di estremi a,x contenuto nell'aperto* Ω.

Allora per ogni $u \in S_1^*$ *esiste almeno un punto* $\xi \in \sum$ *tale che si abbia* :

$$< u \, , \, F(x)-F(a) > \; = \; < u \, , \, dF(\xi)(x-a) > .$$

Si ha altresi':

$$\| F(x) - F(a) \| \leq \sup_{s \in \sum} \| dF(s \| \cdot \| x-a \| \quad , \tag{5}$$

nonche' :

$$\| F(x)-F(a)-dF(a)(x-a) \| \leq \sup_{s \in \sum} \| dF(s)-dF(a) \| \cdot \| x-a \| \quad . \tag{6}$$

Prima della dimostrazione del teorema vogliamo esporre una considerazione sull'enunciato . E' naturale domandarsi se esiste un punto $\xi \in \sum$ tale che si abbia :

$$F(x)-F(a) = dF(\xi)(x-a) \tag{7}$$

conformemente a quanto si verifica per una funzione reale di una variabile reale, dato che la prima affermazione del teorema potrebbe chiamarsi teorema di Lagrange *in forma debole*. Orbene la risposta e' negativa., e per provare cio' basta osservare ad esempio che per la funzione di [0,1] in $\mathbf{R}^2$ definita ponendo

$$F(x) = (x^3-x^2 \, , \, x^3-2x^2)$$

non esiste alcun punto $\xi \in [0,1]$ tale da risultare F(1)-F(0) = F'(ξ) ; infatti F'(x) , ossia ($3x^2$-2x , $3x^2$-4x) , non coincide per alcun x con F(1)-F(0).che e' la coppia (0,-1).

L'estensione della (7) alle funzioni tra spazi di Banach e' una disuguaglianza, appunto la (5). che fornisce una maggiorazione della norma $\| F(x) - F(a) \|$ mediante la norma $\| x-a \|$. Abbiamo aggiunto la (6), che come vedremo e' una conseguenza della (5). perche' da essa si dedurra' la condizione sufficiente per la differenziabilita' in senso forte dianzi accennata.

DIM. (2.4). Fissati $u \in S_1^*$ ed $y \in S_1$ consideriamo la funzione :

$$f : t \in [0,1] \to < u \, , \, F(a+ty) > .$$

Per ogni $t \in [0,1]$ si ha:

$$\frac{f(t+h)-f(t)}{h} = < u \, , \, \frac{F(a+ty+hy)-F(a+ty)}{h} >$$

e quindi :

$$f'(t) = < u \, , \, dF(a+ty)(y) > .$$

D'altra parte,per il teorema di Lagrange esiste $\theta \in]0,1[$ tale che si abbia

$$f(1)-f(0)=f'(\theta) \ ,$$

ossia :

$$< u \ , F(a+y)-F(a) > \ = \ < u \ , dF(a+\theta y)(y) > \ .$$

Posto $\xi = a + \theta y$ e scelto $y=x-a$, da tale relazione segue quella che precede la (5).

La (5) e' banale se $F(x)=F(a)$. Nel caso $F(x) \neq F(a)$,in virtu' di un corollario del teorema di Hahn-Banach [cfr.)9.3) pag.184], posto $z=F(x)-F(a)$ esiste $u \in S_1^*$ tale da risultare $| <u \ , z> | = \| u \| \cdot \| z \|$ e $\| u \| = 1$,ossia :

$$(8) \qquad | <u \ , F(x)-F(a)> | \ = \ \| F(x) \quad F(a) \| \ .$$

In virtu' della prima affermazione del teorema, gia' acquisita, tenuto conto della (8) esiste $\xi \in \sum$ tale che si abbia :

$$(9) \qquad \| F(x) \quad F(a) \| = | \ < u \ , dF(\xi)(x-a) > \ | \ \leq \ \| u \| \ \sup_{s \in \sum} \| dF(s) \| \cdot \| x-a \|$$

donde la (5) dato che $\| u \| = 1$.

Allo scopo di stabilire la (6) consideriamo la funzione G definita ponendo :

$$G(x) = F(x) - dF(a)(x-a) \qquad \forall x \in \Omega \ .$$

Qualunque sia $s \in \sum$,per la (2.1) di pag.290 il funzionale $dG(s)$ e' la differenza tra $dF(s)$ e il differenziale della funzione lineare $dF(a)$,dato che il differenziale della funzione costante $x \to dF(a)(a)$ e' nullo [(2.3) pag.291]: dunque $dG(s)=dF(s)-dF(a) \ \forall s \in \sum$. Dopo cio',applicando la (5) alla funzione G si ottiene la (6) .

In base al teorema di Lagrange si dimostra subito il seguente :

(2.5). Siano S, S_1 spazi vettoriali normati, Ω un aperto di S, F un'applicazione di Ω in S_1. Se F e' differenziabile in senso debole in un intorno $I \subseteq \Omega$ del punto $x_0 \in \Omega$, e se la funzione:

$$\Phi : x \in I \longrightarrow dF(x) \in \mathcal{L}(S, S_1)$$

e' continua in x_0, allora F e' differenziabile nel punto x_0 in senso forte.

DIM. Per ogni $y \in S$ tale che il segmento di estremi x_0, x_0+y sia contenuto in I, in virtù della (6) risulta:

$$\frac{\|F(x_0+y)-F(x_0)-dF(x_0)(y)\|}{\|y\|} \leq \sup_{\theta\in[0,1]} \|dF(x_0+\theta y)-dF(x_0)\|.$$

Per la supposta continuità della funzione Φ nel punto x_0, il secondo membro tende a 0 per $y\to 0$: quindi l'operatore $A=dF(x_0)$, lineare e continuo, verifica la (4) di pag. 290.

Dimostriamo infine i teoremi seguenti.

(2.6). Siano S uno spazio vettoriale normato, F un funzionale reale definito in $X \subseteq S$, differenziabile in senso debole in un punto $x_0 \in \overset{\circ}{X}$.

Condizione necessaria affinché x_0 sia punto di massimo o di minimo relativo per F è che risulti $dF(x_0)=0$.

DIM. Detto δ il raggio di un cerchio aperto di centro x_0 contenuto in X, per ogni fissato $y\in S$ sia $t\in R-\{0\}$ tale che $|t|<\delta/\|y\|$. Se ad es. x_0 è di minimo relativo per F, per t appartenente ad un opportuno intorno di 0 risulta $F(x_0+ty)\geq F(x_0)$, e quindi il rapporto:

$$\frac{F(x_0+ty)-F(x_0)}{t}$$

è positivo o negativo secondo che sia $t>0$ o $t<0$. Pertanto il suo limite per $t\to 0$, ossia $\langle dF(x_0), y\rangle$, è nullo. Siccome ciò vale $\forall y\in S$, l'asserto è provato.

(2.7). Siano S uno spazio vettoriale normato, K un sottoinsieme convesso di S, F un funzionale reale definito in K, differenziabile in senso debole in un punto $x_0 \in K$.

Condizione necessaria affinché $F(x_0)$ sia il minimo di F in K è che risulti:

(10) $$\langle dF(x_0), y-x_0\rangle \geq 0 \quad \forall y\in K.$$

Se K è una varietà lineare di S, diciamola x^*+H con $x^*\in S$ ed H sottospazio vettoriale di S, condizione necessaria affinché $F(x_0)$ sia il minimo di F in K è che risulti:

(11) $$\langle dF(x_0), h\rangle = 0 \quad \forall h\in H.$$

DIM. Fissato $y\in K$, per ogni $t\in[0,1]$ il punto $x_0+t(y-x_0)$ appartiene a K, e quindi risulta:

$$F(x_0+t(y-x_0)) \geq F(x_0),$$

nonché, per $t\in\,]0,1]$:

$$\frac{F(x_0+t(y-x_0))-F(x_0)}{t} \geq 0.$$

Passando al limite per $t\to 0$ ne consegue la (10).

Supponiamo ora che sia $K=x^*+H$; essendo x_0 un punto di K, esiste $h_0\in H$ tale che $x_0=x^*+h_0$. Qualunque sia $h\in H$, siccome il punto $y=x_0+h$ $(=x^*+h_0+h)$ appartiene a K, per la (10) risulta:

$$\langle dF(x_0), h\rangle \geq 0$$

e con lo stesso discorso si deduce che $\langle dF(x_0), -h\rangle \geq 0$, ossia $\langle dF(x_0), h\rangle \leq 0$.
Pertanto sussiste la (11) e l'assunto è dimostrato.

Esempio 1. Differenziale di una funzione di R^n in R^m.

Sia f una funzione vettoriale di n variabili reali, $f=(f_1,\dots,f_m)$, definita in un aperto Ω. Le componenti di f siano differenziabili in Ω (in senso classico). Fissato $x=(x_1,\dots,x_n)\in\Omega$, per ogni $y=(y_1,\dots,y_n)\in R^n$ la funzione della variabile reale t:

$$\Phi(t)=\big(f_1(x+ty), f_2(x+ty,\dots,f_m(x+ty)\big)$$

è definita in un intorno di 0 ed è derivabile nel punto $t=0$. La derivata nel punto 0 della coordinata r^{ma} di $\Phi(t)$, cioè della funzione

$$f_r(x+ty)=f_r(x_1+ty_1, x_2+ty_2,\dots,x_n+ty_n)$$

ha l'espressione:

$$(12)\qquad \frac{\partial f_r}{\partial x_1}(x)\,y_1+\frac{\partial f_r}{\partial x_2}(x)\,y_2+\cdots+\frac{\partial f_r}{\partial x_n}(x)\,y_n$$

e quindi $\Phi'(0)$ è il vettore di R^m le cui componenti sono date dalla (12) con $r=1,2,\dots,m$. Tale vettore è il prodotto della matrice jacobiana:

$$(13)\qquad A(x)=\begin{pmatrix} \frac{\partial f_1}{\partial x_1}(x), \dots, \frac{\partial f_1}{\partial x_n}(x) \\ \dots\dots\dots\dots \\ \frac{\partial f_m}{\partial x_1}(x), \dots, \frac{\partial f_m}{\partial x_n}(x)\end{pmatrix}$$

per il vettore colonna y di componenti $y_1,\dots,y_n$, sicché risulta:

$$\Phi'(0)=A(x)y \qquad \forall y\in R^n.$$

Pertanto [cfr. (1) pag. 288] f è differenziabile secondo Gateaux in ciascun punto x di Ω, e si ha:

$$\langle df(x), y\rangle = A(x)y \qquad \forall y\in R^n.$$

Si può dire allora che per ogni $x\in\Omega$ il differenziale $df(x)$ è la matrice jacobiana (13) (identificata con l'operatore lineare di R^n in R^m che essa rappresenta).

Si verifica facilmente che in ciascun punto x di Ω la f è differenziabile anche in senso forte. Peraltro ciò consegue dal teor. (2.5) di pag. 293 se si suppone che le derivate parziali di f siano continue in Ω.

Esempio 2. L'integrale del calcolo delle variazioni e l'equazione di Eulero.

Per dare un esempio sulla utilizzazione dei teoremi (2.6),(2.7) , ci occuperemo qui del cosiddetto *integrale del calcolo delle variazioni*, che rappresenta un funzionale da minimizzare: la ricerca del minimo di un funzionale, cui si riconducono diverse questioni di ottimizzazione, costituisce il problema fondamentale di un importante settore dell'Analisi Funzionale denominato appunto *Calcolo delle variazioni.*

Prima di esporre l'argomento, ci sembra utile illustrare con qualche esempio il tipo di problema che conduce alla minimizzazione di un funzionale, cominciando con i casi più semplici in cui il funzionale è una funzione di più variabili reali.

1. Assegnati un piano α e un punto P dello spazio, determinare il punto P_0 di α che ha la minima distanza da P.

2. Assegnati un sottoinsieme S di $\mathbf{R}^n$ e un punto P di $\mathbf{R}^n$, determinare un punto P_0 di S (se esiste) che abbia la minima distanza da P. Il problema è una generalizzazione di quello indicato al punto 1.

3. Assegnato un numero positivo l , sia S l'insieme costituito dalle curve piane Γ generalmente regolari, semplici e chiuse, aventi lunghezza l. Ogni $\Gamma \in S$ è la frontiera di un dominio limitato e misurabile $D(\Gamma)$: determinare Γ in modo che $D(\Gamma)$ abbia la massima area. Si tratta di "massimizzare" il funzionale $\Phi : \Gamma \in S \to area\ D(\Gamma)$, cioè di minimizzare il funzionale $-\Phi$.

4. Assegnati due punti A e B di una superficie S, determinare la curva tracciata sulla superficie , congiungente A e B , rettificabile ed avente la minima lunghezza (*curva geodetica*). Detto Λ l'insieme delle curve in questione, si tratta di minimizzare il funzionale $\Phi : \Gamma \in \Lambda \to l(\Gamma)$, essendo $l(\Gamma)$ la lunghezza di Γ .

5. Assegnato un numero positivo σ , sia S l'insieme costituito dalle parti misurabili di $\mathbf{R}^3$ la cui frontiera è una superficie misurabile di area σ : determinare l'elemento di S che ha il massimo volume .Si tratta di "massimizzare" il funzionale $\Phi : A \in S \to vol.A$, cioè di minimizzare il funzionale $-\Phi$.

6. Siano A e B due punti di un piano verticale α , essendo A situato "più in alto" di B. Pensiamo ad una massa puntiforme P, soggetta alla forza peso, che si muova da A a B con velocità iniziale nulla, su una guida giacente sul piano α ed avente la forma di una curva Γ . Il tempo che P impiega per effettuare il movimento descritto dipende dalla traiettoria: determinare la curva Γ in modo che P impieghi il minor tempo possibile (*problema della brachistocrona*). Denotando con S l'insieme costituito dalle curve piane Γ congiungenti A e B , giacenti su α , e con $T(\Gamma)$ il tempo impiegato dalla massa P per percorrere la traiettoria Γ , si tratta di minimizzare il funzionale $\Phi : \Gamma \in S \to T(\Gamma)$.

7. In uno spazio di Hilbert siano dati un punto P e un sottoinsieme K non vuoto, chiuso e convesso. Determinare il punto di K che ha la minima distanza da P.

Il problema , che è una generalizzazione di quello indicato al punto 2, equivale a quello di minimizzare il funzionale (6) di pag. 116.

8. In opportune ipotesi la disequazione variazionale (4) di pag.255 , a cui si possono ricondurre svariati problemi, equivale alla minimizzazione del funzionale (10) di pag. 256.

* * *

Sia $\varphi = \varphi(x, u_0, u_1)$ una funzione reale di tre variabili reali definita in $\bar{\Omega} \times R \times R$, essendo $\Omega =]a,b[$ un intervallo di R.

Supporremo che φ sia dotata di derivate parziali prime e seconde, continue e limitate (1).

Consideriamo il funzionale:

$$(14) \qquad F : u \in C^1(\bar{\Omega}) \longrightarrow \int_a^b \varphi(x, u(x), u'(x))\, dx$$

che è rappresentato dal cosiddetto integrale del calcolo delle variazioni.

Fissato $u \in C^1(\bar{\Omega})$ e posto, per ogni $v \in C^1(\bar{\Omega})$:

$$\Phi(t) = F(u + tv) \quad \forall t \in R,$$

la funzione $\Phi(t)$ è derivabile nel punto $t=0$ e risulta:

$$(15) \qquad \Phi'(0) = \int_a^b \left[\frac{\partial \varphi}{\partial u} v + \frac{\partial \varphi}{\partial u'} v' \right] dx$$

dove, com'è consuetudine, abbiamo indicato con $\partial\varphi/\partial u$, $\partial\varphi/\partial u'$ le derivate parziali di $\varphi(x, u_0, u_1)$ rispetto alla seconda e alla terza variabile calcolate nel punto $(x, u(x), u'(x))$.

La corrispondenza che ad ogni $v \in C^1(\bar{\Omega})$ associa il secondo membro della (15) è un funzionale lineare, ed è ovviamente continuo dato che le derivate parziali di φ sono per ipotesi limitate. Pertanto [cfr. (1) pag. 288] il funzionale (14) è differenziabile secondo Gateaux in ciascun punto $u \in C^1(\bar{\Omega})$, e risulta:

$$(16) \qquad \langle dF(u), v \rangle = \int_a^b \left[\frac{\partial \varphi}{\partial u} v + \frac{\partial \varphi}{\partial u'} v' \right] dx \qquad \forall v \in C^1(\bar{\Omega}).$$

Mostriamo che per ogni $u \in C^1(\bar{\Omega})$ il funzionale F è differenziabile nel punto u anche in senso forte. Proveremo che sussiste la (4) di pag. 290 con $x_0 = u$ ed $A = dF(u)$.

DIM. Ricordando le (14), (16), la differenza $F(u+v) - F(u) - dF(u)\,v$ è l'integrale esteso all'intervallo $[a,b]$ della funzione:

$$G(v) = \left[\varphi(x, u+v, u'+v') - \varphi(x, u, u') \right] - \frac{\partial \varphi}{\partial u}(x, u, u')\, v - \frac{\partial \varphi}{\partial u'}(x, u, u')\, v'$$

dove per brevità abbiamo scritto u, v, u', v' in luogo di $u(x), v(x), u'(x), v'(x)$.

---

(1) Per semplicità abbiamo formulato ipotesi sovrabbondanti.

Posto $\psi(t)=\varphi(x,u(x)+tv(x),u'(x)+tv'(x))$, nel secondo membro della precedente uguaglianza la differenza tra parentesi quadre è uguale a $\psi(1)-\psi(0)$, e quindi si può esprimere come l'integrale di $\psi'(t)$ da 0 ad 1; ciascuno degli altri addendi si può esprimere come l'integrale rispetto a t dell'addendo stesso. Calcolando poi la derivata $\psi'(t)$, si trova che la funzione $G(v)$ può essere espressa come segue:

$$G(v)=\int_0^1\left[v\left(\frac{\partial\varphi}{\partial u}(x,u+tv,u'+tv')-\frac{\partial\varphi}{\partial u}(x,u,u')\right)+v'\left(\frac{\partial\varphi}{\partial u'}(x,u+tv,u'+tv')-\frac{\partial\varphi}{\partial u'}(x,u,v)\right)\right]dt.$$

Applicando alle funzioni $\partial\varphi/\partial u$ e $\partial\varphi/\partial u'$ il teorema di Lagrange per le funzioni reali di tre variabili reali, è chiaro che, detto ℓ un numero maggiorante dei moduli delle derivate seconde di φ, risulta:

$$|G(v)|\le\ell\int_0^1(v^2+v'^2)t\,dt=\frac{\ell}{2}(v^2+v'^2)\le\ell(|v|+|v'|)^2\le\ell\|v\|^2_{C^1(\bar\Omega)}.$$

Ricordando il significato di $G(v)$, ne segue che:

$$\frac{|F(u+v)-F(u)-dF(u)v|}{\|v\|}\le\ell(b-a)\|v\|$$

sicché il primo membro tende a 0 per $\|v\|\to0$. L'assunto è così dimostrato.

La differenziabilità in senso debole del funzionale (14) consente di applicare i teoremi (2.6), (2.7) di pag. 294, al fine di determinare i punti di $C^1(\bar\Omega)$ tra i quali sono compresi gli eventuali punti che minimizzano il funzionale.

Il minimo è da ricercare, a seconda dei casi, o in tutto lo spazio $C^1(\bar\Omega)$, o nel sottospazio:

$$H=\{h\in C^1(\bar\Omega):\ h(a)=h(b)=0\},\tag{17}$$

o più in generale, fissati $\alpha,\beta\in R$, nel sottoinsieme di $C^1(\bar\Omega)$:

$$\{v\in C^1(\bar\Omega):\ v(a)=\alpha\ ,\ v(b)=\beta\}\tag{18}$$

che è la varietà lineare u^*+H di $C^1(\bar\Omega)$, essendo u^* un elemento di $C^1(\bar\Omega)$ tale che $u^*(a)=\alpha$, $u^*(b)=\beta$ (scelto arbitrariamente).

In virtù del teor. (2.6) di pag. 294, affinché $u\in C^1(\bar\Omega)$ minimizzi F è necessario che per ogni $v\in C^1(\bar\Omega)$ si abbia [cfr. (16)]:

$$\int_a^b\left[\frac{\partial\varphi}{\partial u}v+\frac{\partial\varphi}{\partial u'}v'\right]dx=0\tag{19}$$

dove le derivate parziali di φ si intendono calcolate nel punto $(x,u(x),u'(x))$.

Affinché u minimizzi F non in tutto $C^1(\bar\Omega)$, ma nel sottospazio (17) o nella varietà lineare (18) (ossia in u^*+H), per il teor. (2.7) di pag. (294) è necessario che la (19) sussista per ogni $v\in H$ (nei due casi, $u\in H$ o $u\in u^*+H$).

Si dimostra che se, per ogni v come dianzi specificato, è verificata la condizione (19), allora la funzione:

$$g(x) = \frac{\partial \varphi}{\partial u'}(x, u(x), u'(x))$$

è assolutamente continua, quindi derivabile quasi ovunque, nell'intervallo $[a,b]$.

Da ciò si deduce, come andiamo a dimostrare, che:

Se u minimizza il funzionale F in tutto lo spazio $C^1(\bar{\Omega})$, o nel sottospazio (17) o nella varietà lineare (18), allora necessariamente u è soluzione dell'equazione differenziale:

(20) $$\frac{\partial}{\partial x}\frac{\partial \varphi}{\partial u'}(x, u(x), u'(x)) = \frac{\partial \varphi}{\partial u}(x, u(x), u'(x)) .$$

La (20) è nota come equazione di Eulero, o di Eulero-Lagrange.

DIM. Per l'ipotesi posta la (19) sussiste in particolare per ogni $v \in C_0^\infty(\Omega)$; decomposto l'integrale nella somma di due integrali, nel secondo di essi è lecito effettuare l'integrazione per parti col fattore differenziale $v(x)dx$. Essendo $v(a) = v(b) = 0$, dalla (19) si deduce allora:

$$\int_a^b v\left[\frac{\partial \varphi}{\partial u} - \frac{d}{dx}\frac{\partial \varphi}{\partial u'}\right]dx = 0 \qquad \forall v \in C_0^\infty(\Omega)$$

e quindi u verifica la (20) q.o. in $[a,b]$ (cfr. (11) pag. 150).

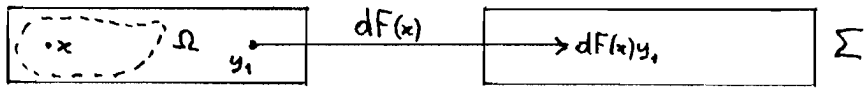
A pag. 302 riporteremo due esempi sul funzionale (14).

3. Differenziali di ordine superiore.

Consideriamo la funzione:

$$F : \Omega \subseteq S \to \Sigma$$

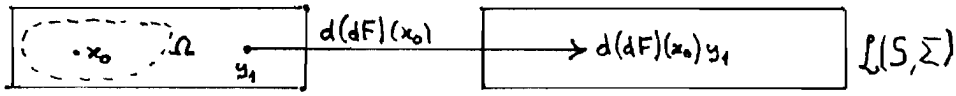
con S, Σ spazi di Banach ed Ω insieme aperto. La F sia differenziabile in ogni punto $x \in \Omega$, in senso forte [pag. 288 e 290]. Ricordiamo che $dF(x)$ è, per ogni $x \in \Omega$, un operatore lineare e continuo, definito in tutto lo spazio S e a valori in Σ: perciò per ogni $y_1 \in S$ il suo corrispondente mediante $dF(x)$, ossia $dF(x)y_1$, è un elemento di Σ.



Fissato un punto $x_0 \in \Omega$, consideriamo ora la funzione definita in Ω:

(1) $$dF : x \in \Omega \to dF(x) \in \mathcal{L}(S, \Sigma)$$

che è a valori nello spazio degli operatori lineari e continui di S in Σ (pag. 203). Se tale funzione è differenziabile nel punto x_0 in senso forte, si dice che la F è differenziabile due volte nel punto x_0: il differenziale in x_0 della funzione (1), ossia $d(dF)(x_0)$, è allora un operatore lineare e continuo, definito in S ed a valori in $\mathcal{L}(S, \Sigma)$.



Dunque per ogni $y_1 \in S$ il suo corrispondente mediante $d(dF)(x_0)$, ossia $d(dF)(x_0)y_1$, è a sua volta un operatore lineare e continuo di S in Σ: il corrispondente mediante quest'ultimo di un punto $y_2 \in S$ è allora un elemento di Σ: orbene, la corrispondenza che ad $(y_1, y_2) \in S \times S$ associa tale elemento di Σ si denota con $d^2F(x_0)$ e si chiama il differenziale secondo, o di ordine 2, della funzione F nel punto x_0. Notiamo che

la corrispondenza $d^2F(x_0)$ non coincide con $d(dF)(x_0)$: la prima è un'applicazione bilineare e continua di $S\times S$ in Σ (1), mentre la seconda, come già detto, è un operatore lineare e continuo di S in $\mathcal{L}(S,\Sigma)$. Peraltro la norma di $d^2F(x_0)$, nello spazio $\mathcal{B}(S\times S,\Sigma)$ delle applicazioni bilineari e continue di $S\times S$ in Σ [cfr. (7) pag. 211], è uguale alla norma di $d(dF)(x_0)$, nello spazio degli operatori lineari e continui di S in $\mathcal{L}(S,\Sigma)$ [cfr. (2') pag. 203], dato che i due spazi sono equivalenti [(2.2) pag. 211].

Per ogni $y\in S$, l'elemento $d^2F(x_0)(y,y)$ dello spazio Σ si denota col simbolo:

$$d^2F(x_0)[y]^2 \quad (2) .$$

Supposto che F sia differenziabile due volte in ciascun punto di Ω, ha significato la funzione:

$$(2) \qquad d^2F : x\in\Omega \to d^2F(x)\in\mathcal{B}(S\times S,\Sigma) .$$

Fissato un punto $x_0\in\Omega$, a partire dalla (2) si introduce, come si è fatto dianzi a partire dalla (1), la locuzione F differenziabile tre volte in x_0, e successivamente si definisce il differenziale di ordine 3 della funzione F nel punto x_0, che è un'applicazione 3-lineare (con ovvio significato del termine) e continua, di S^3 in Σ.

Per ogni $y\in S$, l'elemento $d^3F(x_0)(y,y,y)$ si denota col simbolo $d^3F(x_0)[y]^3$.

Procedendo induttivamente, per ogni $n\in N$ si definisce il differenziale n-mo, o di ordine n, della funzione F nel punto x_0, denotato con $d^nF(x_0)$: è un'applicazione n-lineare (con ovvio significato del termine) e continua, di S^n in Σ. Naturalmente per uniformità si pone $d^1F(x_0)=dF(x_0)$.

Per ogni $y\in S$, l'elemento $d^nF(x_0)(y_1,\dots,y_n)$ di Σ con $y_1=y_2=\dots=y_n=y$ si denota col simbolo $d^nF(x_0)[y]^n$.

Il teorema seguente, che ci limitiamo ad enunciare, fornisce una

---

(1) La continuità si verifica facilmente in base alla (2.1) di pag. 209. Si ha infatti:

$$\|d^2F(x_0)(y_1,y_2)\|_\Sigma \le \|d(dF)(x_0)y_1\|_{\mathcal{L}(S,\Sigma)}\cdot\|y_2\|_S \le \|d(dF)(x_0)\|\cdot\|y_1\|_S\cdot\|y_2\|_S \quad \forall (y_1,y_2)\in S\times S$$

dove la norma di $d(dF)(x_0)$ è intesa nello spazio degli operatori lineari e continui di S in $\mathcal{L}(S,\Sigma)$ [cfr. (3) pag. 203].

(2) Fissati $x_0\in\Omega$ e $y_1,y_2\in S$, per calcolare formalmente $d^2F(x_0)(y_1,y_2)$, e poi $d^2F(x_0)[y]^2$ per $y_1=y_2=y$, occorre utilizzare due volte la nota (1) a piè di pag. 288.

Introdotti due parametri reali t e τ, calcolando nel punto $t=0$ la derivata della funzione:

$$t\to F(x_0+ty_1+\tau y_2)\in\Sigma$$

si ottiene il valore nel punto y_1 dell'operatore lineare $dF(x_0+\tau y_2)$. Calcolando poi la derivata nel punto $\tau=0$ della funzione:

$$\tau\to dF(x_0+\tau y_2)y_1\in\Sigma$$

si ottiene il valore nel punto y_2 dell'operatore lineare $d(dF)(x_0)y_1$, che è appunto $d^2F(x_0)(y_1,y_2)$. Se $y_1=y_2=y$ si perviene a $d^2F(x_0)[y]^2$.

estensione della formula di Taylor.

(3.1). Si supponga che la funzione:

$$F : \Omega \subseteq S \to \Sigma$$

con S, Σ spazi di Banach, sia differenziabile n volte in ogni punto dell'aperto Ω, e che la funzione:

$$d^n F : x \in \Omega \to d^n F(x)$$

sia continua. Allora qualunque siano $x_0 \in \Omega$ ed $h \in S$ tali che il segmento di estremi x_0, x_0+h sia contenuto in Ω, risulta:

$$F(x_0+h) - F(x_0) = dF(x_0)h + \frac{1}{2!} d^2F(x_0)[h]^2 + \cdots + \frac{1}{n!} d^nF(x_0)[h]^n + r_n(h),$$

dove r_n è una funzione a valori in Σ tale da risultare:

$$\lim_{h \to 0} \frac{r_n(h)}{\|h\|^n} = 0.$$

In base alla (3.1) dimostriamo il teorema seguente, che fornisce una condizione sufficiente affinché un punto x_0 sia di minimo relativo per un funzionale reale F.

(3.2). Sia F un funzionale reale definito in un sottoinsieme X di uno spazio di Banach S, differenziabile due volte in un intorno I del punto x_0 interno ad X. La funzione $d^2F : x \in I \to d^2F(x)$ sia continua.
Se risulta $dF(x_0) = 0$, e se la forma bilineare $d^2F(x_0)(y_1, y_2)$ è coercitiva, allora x_0 è punto di minimo relativo per F.

L'ipotesi di coercitività significa [pag.213] che esiste una costante positiva c tale che:

$$d^2F(x_0)[h]^2 \geq c\|h\|^2 \qquad \forall h \in S. \tag{3}$$

DIM. Denotiamo con δ un numero positivo tale:

$$\|h\| < \delta \Rightarrow x_0 + h \in I.$$

Qualunque sia $h \in S$ di norma minore di δ, per la formula di Taylor si ha:

$$F(x_0+h) - F(x_0) = \frac{1}{2} d^2F(x_0)[h]^2 + r_2(h) \tag{4}$$

dove r_2 è una funzione reale tale da risultare:

$$\lim_{h \to 0} \frac{r_2(h)}{\|h\|^2} = 0.$$

Scelto $\delta_1 < \delta$ in modo che si abbia, per $0 < \|h\| < \delta_1$:

$$\frac{\|\tau_2(h)\|}{\|h\|^2} < \frac{c}{4}$$

dalla (4) consegue, tenendo conto della (3):

$$F(x_0+h) - F(x_0) \geq \frac{1}{2} c \|h\|^2 - \frac{1}{4} c \|h\|^2 = \frac{1}{4} c \|h\|^2 > 0 .$$

Dunque x_0 è punto di minimo relativo per F.

---

4. Cenno sull'integrazione di una funzione continua a valori in uno spazio di Banach.

Siano: S uno spazio di Banach, $[a,b]$ un intervallo compatto di R, F un'applicazione di $[a,b]$ in S, continua in $[a,b]$.

Decomponiamo l'intervallo $[a,b]$ in n intervalli mediante i punti:

(1) $$a = t_0 < t_1 < \dots < t_{n-1} < t_n = b .$$

Scelto in ciascun intervallo $[t_i, t_{i+1}]$ un punto ξ_i, consideriamo l'elemento di S:

(2) $$\sum_{i=0}^{n-1} F(\xi_i)(t_{i+1} - t_i) .$$

Con lo stesso procedimento relativo al caso $S = R$ si dimostra che, detta δ la più grande delle ampiezze degli intervalli della decomposizione, esiste il limite per $\delta \to 0$ della somma (2). Esiste cioè un punto y_0 di S, univocamente determinato, godente della seguente proprietà: per ogni $\varepsilon > 0$ esiste $\delta_\varepsilon > 0$ tale che, comunque si decomponga l'intervallo $[a,b]$, mediante i punti (1), in intervalli tutti di ampiezza minore di δ_ε, e comunque si scelga in ciascun intervallo $[t_i, t_{i+1}]$ il punto ξ_i, si abbia:

$$\left\| \sum_{i=0}^{n-1} F(\xi_i)(t_{i+1} - t_i) - y_0 \right\| < \varepsilon .$$

Il limite per $\delta \to 0$ della somma (2) si chiama <u>integrale esteso ad $[a,b]$</u> dell'applicazione F, e si denota con

$$\int_a^b F(t)\,dt .$$

Segnaliamo che risulta:

$$\left\| \int_a^b F(t)\,dt \right\| \leq \int_a^b \|F(t)\|\,dt$$

e che, detto A un operatore lineare e continuo di S in S_1, con S_1 anch'esso spazio di Banach, si ha:

$$\int_a^b AF(t)\,dt = A\int_a^b F(t)\,dt.$$

E' poi evidente che, se y_0 e' un elemento di S e $F(t)=f(t)y_0$, con $f:[a,b]\to\mathbb{R}$ funzione continua, l'integrale di F e' uguale al prodotto di y_0 per l'integrale di f. Cioe' risulta:

$$\int_a^b f(t)y_0\,dt = y_0\int_a^b f(t)\,dt.$$

5. Due esempi sulla minimizzazione del funzionale (14) pag. 296.

1. Si voglia minimizzare il funzionale:

$$F(u)=\int_0^1 u'^2(x)\,dx$$

nell'insieme V costituito dalle funzioni $v\in C^1([0,1])$ tali che $v(0)=0, v(1)=1$.

Con le notazioni di pag. 296, si ha $\varphi(x,u_0,u_1)=u_1^2$ e quindi $\frac{\partial\varphi}{\partial u_0}=0$, $\frac{\partial\varphi}{\partial u_1}=2u_1$. Se esiste $u\in V$ che minimizza F, la funzione u deve essere soluzione dell'equazione di Eulero [cfr. 20 pag. 298], cioe' dell'equazione $\frac{d}{dx}2u'(x)=0$, ossia $u''(x)=0$. Pertanto u deve essere un polinomio del tipo $ax+b$, e siccome deve appartenere a V deve essere $u(x)=x$. Occorre ora stabilire se effettivamente $F(u)$ e' il minimo di F in V. In virtu' della disuguaglianza di Schwarz, per ogni $v\in V$ si ha:

$$1=\int_0^1 v'(x)\,dx=\int_0^1 1\cdot v'(x)\,dx \leq \left(\int_0^1 1^2\,dx\right)^{1/2}\left(\int_0^1 v'^2(x)\,dx\right)^{1/2}=\left(F(v)\right)^{1/2}$$

e quindi $F(v)\geq 1\ \forall v\in V$. Siccome $1=F(u)$, ne segue che u minimizza F.

2. Si voglia minimizzare il funzionale:

$$F(u)=\int_0^1 \left[u'^2(x)+u'^3(x)\right]dx$$

nel sottospazio $H=\{h\in C^1([0,1]) : h(0)=h(1)=0\}$.

L'unica soluzione appartenente ad H dell'equazione di Eulero [ossia $u''(2+6u')=0$] e' la funzione identicamente nulla in $[0,1]$: pero' tale funzione non minimizza F in H, in quanto come andiamo a dimostrare la F non e' limitata inferiormente in H.

Posto $v(x)=x^3-x\ \forall x\in[0,1]$, $F(v)$ e' la somma dell'integrale di v'^2, diciamolo λ, e dell'integrale di v'^3 che e' uguale a $16/35$. Orbene, qualunque sia $c\in\mathbb{R}$ la funzione $h=cv$ appartiene ad H, e siccome risulta

$$F(cv)=\frac{16}{35}c^3+\lambda c^2,$$

quindi $F(cv)\to-\infty$ per $c\to-\infty$, la F non e' limitata inferiormente in H.

Appendice

TEOREMA DI SOBOLEV

Si consideri lo spazio $W^{m,p}(\Omega)$ (m intero positivo, $p \geq 1$), con Ω aperto di R^n dotato della proprietà di cono [1]. E' opportuno evidenziare che se in particolare l'aperto Ω e' di classe $C^{0,1}$, ossia a frontiera localmente lipschitziana [2], qualunque siano m, p risulta $W^{m,p}(\Omega) = H^{m,p}(\Omega)$; tale uguaglianza sussiste inoltre se Ω e' un intervallo di R.

Fissato un intero $r \in [1, m]$, si consideri il numero reale:

(1) $$\frac{1}{p} - \frac{r}{n} \quad .$$

Caso $\frac{1}{p} - \frac{r}{n} > 0$

Posto:

$$\frac{1}{q} = \frac{1}{p} - \frac{r}{n} \quad ,$$

si ha:

$$W^{m,p}(\Omega) \subset W^{m-r,q'}(\Omega) \begin{cases} \text{con immersione continua [3]} & \forall q' \in [p, q] \\ \text{se } \Omega \text{ e' limitato, con immersione compatta [4]} & \forall q' \in [p, q[\text{ e per } q' = q \text{ se } r = m. \end{cases}$$

Caso $\frac{1}{p} - \frac{r}{n} = 0$

Si ha:

$$W^{m,p}(\Omega) \subset W^{m-r,q}(\Omega) \begin{cases} \text{con immersione continua} & \forall q \in [p, +\infty[\text{ e per } q = +\infty \text{ se } p = 1 \\ \text{se } \Omega \text{ e' limitato, con immersione compatta} & \forall q \in [p, +\infty[\quad . \end{cases}$$

---

(1) Cfr. nota (1) pag. 101.

(2) Con ciò si intende che Ω e' l'interno di un dominio e gode della proprietà: per ogni $x \in \partial\Omega$ esiste un intorno aperto I_x di x tale che l'insieme $I_x \cap \partial\Omega$ sia rappresentabile mediante un'equazione del tipo $\xi_n = \varphi(\xi_1, \ldots, \xi_{n-1})$ con φ definita in un aperto di R^{n-1} ed ivi lipschitziana.
Si dimostra che ogni aperto di classe $C^{0,1}$ e' dotato della proprietà di cono.

(3) Cfr. pag. 63. Se X, Y sono spazi vettoriali normati con $X \subseteq Y$, il fatto che l'immersione di X in Y e' continua significa che esiste una costante c tale da risultare:

$$\|u\|_Y \leq c \|u\|_X \quad \forall u \in X.$$

(4) Cfr. pag. 165.

Caso $\frac{1}{p} - \frac{r}{n} < 0$

Si ha:

$$W^{m,p}(\Omega) \subset W^{m-r,q}(\Omega) \quad \text{con immersione compatta } \forall q \in [1, +\infty[$$

nonché:

$$W^{m,p}(\Omega) \subset C^{m-r}(\bar{\Omega}) \quad \text{con immersione compatta} \quad (1).$$

Inoltre, posto:

$$\lambda = r - \frac{n}{p} \ ,$$

ossia detto λ l'opposto del numero reale negativo (1) moltiplicato per n, se $\lambda < 1$ si ha:

$$W^{m,p}(\Omega) \subset C^{m-r,\lambda'}(\bar{\Omega}) \begin{cases} \text{con immersione continua } \forall \lambda' \in]0, \lambda] \\ \text{se } \Omega \text{ è limitato e di classe } C^{0,1} \\ \text{con immersione compatta} \end{cases} \forall \lambda' \in]0, \lambda[$$

e se $\lambda = 1$ si ha:

$$W^{m,p}(\Omega) \subset C^{m-r,\lambda'}(\bar{\Omega}) \begin{cases} \forall \lambda' \in]0,1[\text{ e per } \lambda' = 1 \text{ se } r = n+1 \ , \ p = 1 \\ \text{con immersione continua ; se } \Omega \text{ è limitato} \\ \text{e di classe } C^{0,1}, \text{ con immersione compatta .} \end{cases}$$

Se Ω è convesso l'ultima inclusione (relativa al caso $\lambda = 1$) sussiste anche per $\lambda' = 1$, con immersione continua.

Si noti che se $n = 1$, $p = 2$, $r = 1$ il numero reale (1) è uguale a $-1/2 < 0$, e quindi occorre far riferimento al terzo caso. Risulta $\lambda = \frac{1}{2}$ e si ritrova il teorema enunciato a pag. 105.

---

(1) Ovviamente l'inclusione va intesa nel senso che ogni elemento u di $W^{m,p}(\Omega)$ ha un rappresentante $\tilde{u}$ che appartiene a $C^{m-r}(\bar{\Omega})$; l'immersione, naturalmente, è l'operatore:

$$u \in W^{m,p}(\Omega) \longrightarrow \tilde{u} \in C^{m-r}(\bar{\Omega}) .$$

MISURE SPETTRALI IN UNO SPAZIO DI HILBERT

RAPPRESENTAZIONE SPETTRALE DI UN OPERATORE AUTOAGGIUNTO

Sia S uno spazio di Hilbert di dimensione infinita.

Nel n. 4 di pag. 245 ci siamo occupati della rappresentazione spettrale di un operatore autoaggiunto $A: S \to S$ che sia anche compatto, sicché l'insieme degli autovalori di A è numerabile [(3.2) pag. 244]. Precisamente abbiamo visto che un siffatto operatore A si può esprimere nella forma:

$$A = \sum_{n=1}^{\infty} \lambda_n P_n$$

dove i termini della successione $\{\lambda_n\}$ sono gli elementi non nulli dello spettro puntuale di A, ossia gli autovalori di A diversi da zero, e $\{P_n\}$ è una successione di proiettori associata alla successione $\{\lambda_n\}$ (1).

Se si considera un operatore autoaggiunto A non necessariamente compatto, si può costruire una analoga rappresentazione di A, ma alla successione di proiettori $\{P_n\}$ subentra una famiglia di proiettori $(P_\lambda)_{\lambda \in R}$, e alla serie subentra un integrale. Nelle considerazioni che seguono accenneremo a tale rappresentazione.

Per ogni $\lambda \in R$ sia P_λ un proiettore definito in S, cioè un operatore che ad ogni $x \in S$ associa la proiezione ortogonale di x su un sottospazio chiuso H_λ di S. La famiglia di proiettori:

$$(P_\lambda)_{\lambda \in R}$$

si chiama una <u>misura spettrale</u> quando soddisfa alle seguenti condizioni:

1) E' crescente, ossia tale che $\lambda > \mu \Longrightarrow P_\lambda \geq P_\mu$ (2).

2) E' continua a destra, cioè tale che $\forall \lambda_0 \in R$ si abbia: $\lim_{\lambda \to \lambda_0 +} P_\lambda x = P_{\lambda_0} x \quad \forall x \in S$.

3) Esistono due numeri reali m, M con $m < M$ tali che, detta I l'identità in S, si abbia:

$$P_\lambda = 0 \quad \forall \lambda \leq m \;, \qquad P_\lambda = I \quad \forall \lambda \geq M \;, \qquad 0 < P_\lambda < I \quad \forall \lambda \in]m, M[\;.$$

Si dimostra che per ogni $\lambda_0 \in R$ e per ogni $x \in S$ esiste anche il limite sinistro di $P_\lambda x$ per $\lambda \to \lambda_0^-$ e che l'operatore $P_{\lambda_0 - 0} : x \in S \to \lim_{\lambda \to \lambda_0^-} P_\lambda x$ è anch'esso un proiettore.

Supponiamo ora di aver fissato, una volta per tutte, due punti x, y di S, e poniamo, supposto che $(P_\lambda)_{\lambda \in R}$ sia una misura spettrale:

$$(1) \qquad \varphi(\lambda) = (P_\lambda x, y) \qquad \forall \lambda \in R \,.$$

Si dimostra che la funzione φ è a variazione limitata in ogni intervallo compatto di R [ossia, $Re(\varphi)$ e $Im(\varphi)$ sono funzioni a variazione limitata nell'intervallo]. Pertanto, se f è una funzione complessa definita in R ed ivi continua, scelto un intervallo $[a, b]$ che contenga nel suo interno l'intervallo $[m, M]$ si può considerare l'integrale di Stieltjes [n. 9 pag. 48]:

$$(2) \qquad \int_a^b f(\lambda)\, d\varphi \,.$$

---

(1) Ricordiamo che per ogni sottospazio chiuso H di S il proiettore P_H di S su H [ch. pag. 121] è un operatore lineare e continuo di norma unitaria, ed è anche autoaggiunto [cfr. es. 6 pag 227].

(2) Per il significato della disuguaglianza, cfr. pag. 223.

Siccome φ è identicamente nulla nell'intervallo $[a,m[$ (cfr. pag. 306, formula (1) e condizione 3)], per la (7) di pag. 48.1 si ha:

$$\int_a^m f(\lambda)\,d\varphi = f(m)\varphi(m)$$

e d'altra parte, essendo φ costante nell'intervallo $[M,b]$, per la (6) di pag. 48.1 si ha:

$$\int_M^b f(\lambda)\,d\varphi = 0 .$$

Di conseguenza risulta [(5) pag. 48.1]:

$$\int_a^b f(\lambda)\,d\varphi = \int_a^m f(\lambda)\,d\varphi + \int_m^M f(\lambda)\,d\varphi + \int_M^b f(\lambda)\,d\varphi =$$

$$= f(m)\varphi(m) + \int_m^M f(\lambda)\,d\varphi .$$

Dunque l'integrale (2) non dipende dalla scelta dei punti a,b tali che $[m,M]\subset]a,b[$, sicché esso può denotarsi con $\int_{-\infty}^{+\infty} f(\lambda)\,d\varphi$. Si ha allora:

$$\int_{-\infty}^{+\infty} f(\lambda)\,d\varphi = f(m)\varphi(m) + \int_m^M f(\lambda)\,d\varphi . \tag{4}$$

Siccome l'integrale (4) non dipende dai valori che f assume nei punti esterni all'intervallo $[m,M]$, esso ha significato anche se f è definita solo nell'intervallo $[m,M]$, giacché si può ritenere la f prolungata arbitrariamente in una funzione continua in R. Analogamente, se in un intorno di un punto $\lambda_0 \in]m,M[$ la funzione $\lambda \to P_\lambda$ e quindi anche $\varphi(\lambda)$ è costante, tenuto conto della (6) di pag. 48.1 è chiaro che l'integrale (4) non si altera modificando la f in un intervallo $[c,d]$ contenuto nell'intorno e tale che $c<\lambda_0<d$: possiamo allora concludere che l'integrale (4) ha senso anche se nel punto λ_0 la funzione f non è continua o non è definita.

Tutto ciò premesso, mostriamo che fissata una funzione reale o complessa f definita in R, ivi continua, ogni misura spettrale $(P_\lambda)_{\lambda\in R}$ individua un operatore $B: S\to S$ lineare e continuo.

Si tratta dell'operatore B che ad ogni $x\in S$ associa l'elemento $\tilde{x}$ di S determinato come segue.

Per ogni $y\in S$ resta determinata la funzione $\varphi(\lambda)$ definita dalla (1) di pag. 306 e quindi resta determinato l'integrale (4); esplicitando in (4) la $\varphi(\lambda)$, consideriamo il funzionale:

$$F_x : y\in S \to \int_{-\infty}^{+\infty} f(\lambda)\,d(P_\lambda x, y)$$

che è un elemento dello spazio S^* duale di S [omettiamo la dimostrazione].

Utilizziamo ora il teorema di Riesz [(7.1) pag. 177], precisamente il comma successivo alla (2) con $u = F_x$, indicando poi con $\tilde{x}$ l'elemento di S ivi indicato con x: dunque per il teorema di Riesz esiste uno ed un solo $\tilde{x}\in S$ tale che:

$$(\tilde{x},y) = \int_{-\infty}^{+\infty} f(\lambda)\,d(P_\lambda x, y) \quad \forall y\in S . \tag{5}$$

Orbene, si dimostra che l'operatore:

$$B : x\in S \to \tilde{x}\in S$$

e' lineare e continuo. Essendo $Bx=\tilde{x}$, la (5) si scrive:

$$(6) \qquad (Bx,y)=\int_{-\infty}^{+\infty} f(\lambda)\,d(P_\lambda x,y) \qquad \forall x,y\in S\ .$$

Per evidenziare che l'operatore B e' determinato dalla funzione f e dalla misura spettrale $(P_\lambda)_{\lambda\in R}$, si scrive:

$$B=\int_{-\infty}^{+\infty} f(\lambda)\,dP_\lambda$$

giacche' tale uguaglianza richiama formalmente la (6) che caratterizza l'operatore B.

Si dimostra che:

I. Se $f(\lambda)=1$, qualunque sia la misura spettrale $(P_\lambda)_{\lambda\in R}$ l'operatore B e' l'identita' in S. Pertanto si ha:

$$(7) \qquad \int_{-\infty}^{+\infty} d(P_\lambda x,y)=(x,y) \qquad \forall (x,y)\in S\times S\ .$$

II. L'aggiunta hermitiano dell'operatore B determinato da f e $(P_\lambda)_{\lambda\in R}$ e' l'operatore B^* determinato dalla funzione coniugata di f e da $(P_\lambda)_{\lambda\in R}$.
Di conseguenza se f e' reale l'operatore B e' autoaggiunto.

* * *

Fissata la funzione f definita ponendo $f(\lambda)=\lambda\ \forall\lambda\in R$, ogni misura spettrale $(P_\lambda)_{\lambda\in R}$ determina l'operatore lineare e continuo di S in S:

$$B=\int_{-\infty}^{+\infty}\lambda\,dP_\lambda\ .$$

che e' autoaggiunto per la proposizione II.

Orbene si dimostra che viceversa, per ogni operatore lineare, continuo e autoaggiunto A di S in S, esiste una e una sola misura spettrale $(P_\lambda)_{\lambda\in R}$ tale che si abbia:

$$(8) \qquad A=\int_{-\infty}^{+\infty}\lambda\,dP_\lambda\ .$$

La suddetta $(P_\lambda)_{\lambda\in R}$ si chiama la misura spettrale associata all'operatore A, e l'uguaglianza (8) si chiama la rappresentazione spettrale di A. Se la misura spettrale che figura in (7) e' quella associata all'operatore A, l'uguaglianza:

$$I=\int_{-\infty}^{+\infty} dP_\lambda$$

si chiama la rappresentazione spettrale, associata all'operatore A, dell'operatore identico $I:S\to S$.

Per il teor. (3.1) di pag. 242, i numeri complessi con parte immaginaria non nulla appartengono tutti all'insieme risolvente di A, e lo spettro residuo di A e' vuoto. Pertanto ciascun numero reale λ_0 o appartiene all'insieme risolvente di A, o allo spettro puntuale, o allo spettro continuo. Ciascuna di tali eventualita' e' caratterizzata da una proprieta' della misura spettrale associata ad A, precisata nel teorema seguente.

III. Se $(P_\lambda)_{\lambda\in\mathbb{R}}$ è la misura spettrale associata all'operatore lineare, continuo e autoaggiunto $A: S \to S$, i numeri m, M menzionati nella condizione 3) di pag. 306 sono rispettivamente gli estremi $m(A), M(A)$ dell'operatore A [pag. 223], e quindi sono il minimo e il massimo dello spettro di A [teor. (3.1) pag. 842].

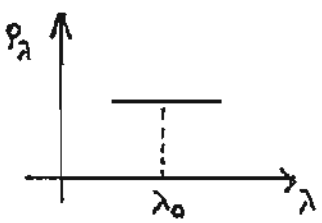
Per ogni $\lambda_0 \in \mathbb{R}$ si ha quanto segue:

1) λ_0 appartiene all'insieme risolvente di A se e solo se esiste un intorno di λ_0 in cui la funzione $\lambda \to P_\lambda$ è costante.

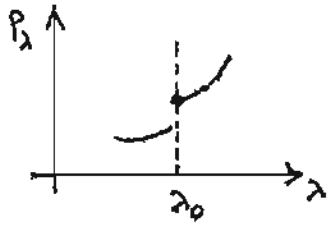
2) λ_0 appartiene allo spettro puntuale di A (ossia è autovalore di A) se e solo se la funzione $\lambda \to P_\lambda$ è discontinua nel punto λ_0 (a sinistra). In tal caso l'operatore $P_{\lambda_0} - P_{\lambda_0 - 0}$ [1] è il proiettore dello spazio S sull'autospazio $\ker(A - \lambda_0 I)$ corrispondente all'autovalore λ_0.

3) λ_0 appartiene allo spettro continuo di A se e solo se la funzione $\lambda \to P_\lambda$ è continua nel punto λ_0 e non esiste alcun intorno di λ_0 in cui è costante.

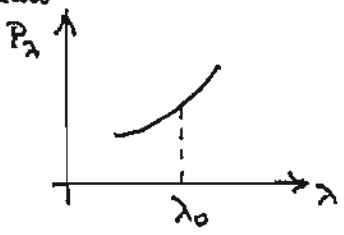
Per comodità del lettore, ci sembra utile schematizzare la funzione $\lambda \to P_\lambda$ come segue, nelle tre eventualità del teorema appena enunciato.



λ_0 appartiene all'insieme risolvente di A



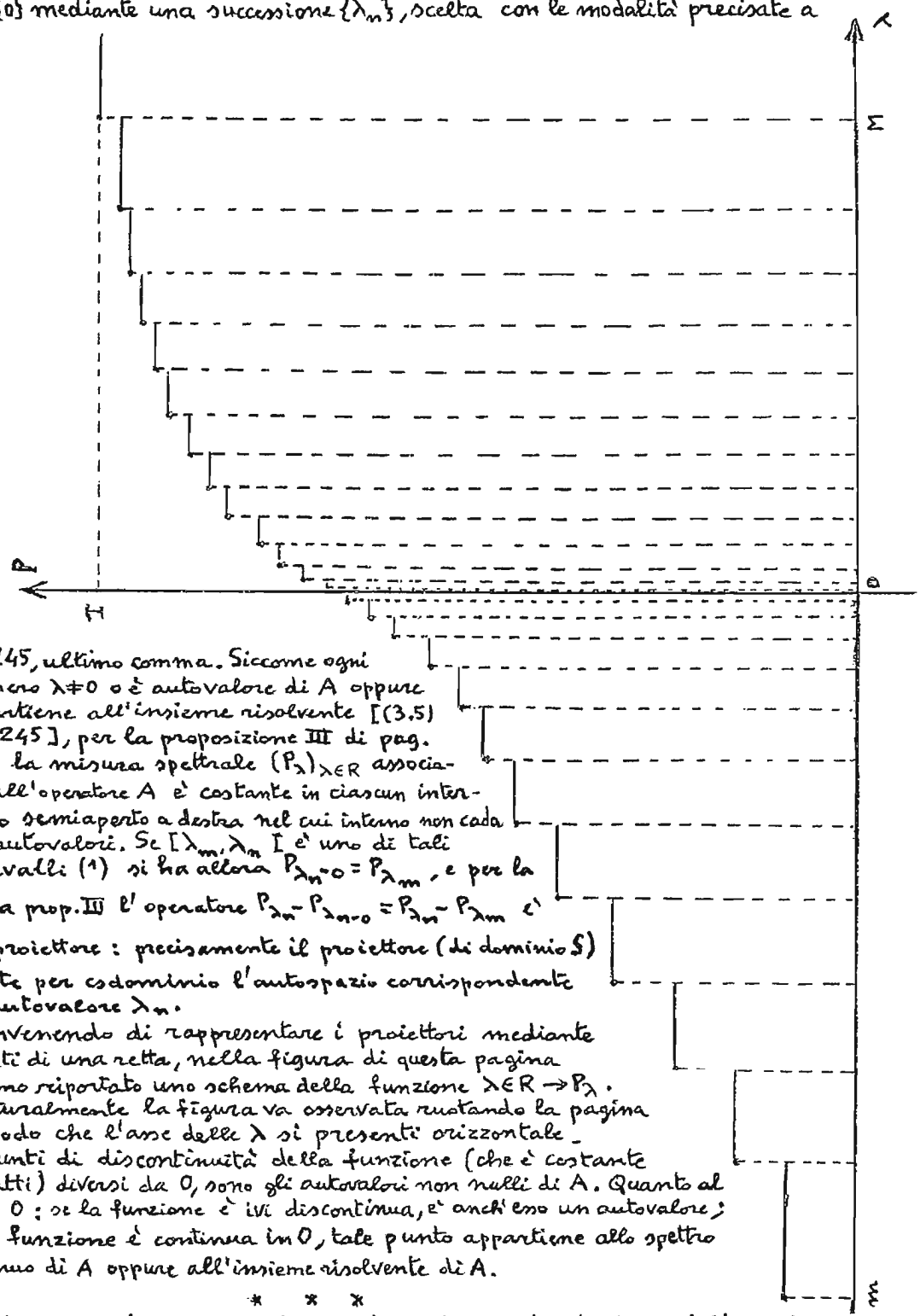
λ_0 è autovalore di A



λ_0 appartiene allo spettro continuo di A

[1] Si tenga presente quanto asserito dopo la condizione 3) di pag. 306.

Se l'operatore lineare e continuo A è compatto, oltre che autoaggiunto, per la (3.2) di pag. 244 l'insieme Λ degli autovalori di A [tutti reali, (3.1) pag.242] è finito o numerabile. Supposto che esso sia numerabile, rappresentiamo $\Lambda - \{0\}$ mediante una successione $\{\lambda_n\}$, scelta con le modalità precisate a pag.245, ultimo comma. Siccome ogni numero $\lambda \neq 0$ o è autovalore di A oppure appartiene all'insieme risolvente [(3.5) pag.245], per la proposizione III di pag. 308.1 la misura spettrale $(P_\lambda)_{\lambda \in R}$ associata all'operatore A è costante in ciascun intervallo semiaperto a destra nel cui interno non cadano autovalori. Se $[\lambda_m, \lambda_n[$ è uno di tali intervalli (1) si ha allora $P_{\lambda_n - 0} = P_{\lambda_m}$, e per la stessa prop. III l'operatore $P_{\lambda_n} - P_{\lambda_n - 0} = P_{\lambda_n} - P_{\lambda_m}$ è un proiettore: precisamente il proiettore (di dominio S) avente per codominio l'autospazio corrispondente all'autovalore λ_n.



Convenendo di rappresentare i proiettori mediante i punti di una retta, nella figura di questa pagina abbiamo riportato uno schema della funzione $\lambda \in R \to P_\lambda$.

Naturalmente la figura va osservata ruotando la pagina in modo che l'asse delle λ si presenti orizzontale.

I punti di discontinuità della funzione (che è costante a tratti) diversi da 0, sono gli autovalori non nulli di A. Quanto al punto 0: se la funzione è ivi discontinua, è anch'esso un autovalore; se la funzione è continua in 0, tale punto appartiene allo spettro continuo di A oppure all'insieme risolvente di A.

* * *

Vogliamo mostrare che nel caso dianzi considerato, in cui l'operatore lineare e continuo A è compatto, oltre che autoaggiunto, la rappresentazione spettrale di A [(8) pag.308] si riduce alla decomposizione spettrale di A introdotta a pag. 247 [(16) pag.246], cioè all'uguaglianza:

(1) Ricordiamo che è decrescente la successione dei moduli $(|\lambda_n|)_{n \in N}$, non la successione $(\lambda_n)_{n \in N}$: perciò non è detto che gli intervalli in questione siano del tipo $[\lambda_{n+1}, \lambda_n[$.

$$A = \sum_{n=1}^{\infty} \lambda_n P_n$$

dove P_n è il proiettore di S sul sottospazio generato da x_n, essendo x_n l'autovettore associato a λ_n con le modalità precisate a pag. 246.

Per semplicità supporremo che gli autovalori di A siano tutti positivi e di molteplicità 1. Ciò comporta che:

1) La successione $\{\lambda_n\}$ è decrescente in senso stretto, oltre che infinitesima, sicchè gli intervalli $[\lambda_m, \lambda_n[$ considerati a pag. 309 sono del tipo $[\lambda_{m+1}, \lambda_n[$.

2) Essendo λ_n di molteplicità 1, l'autospazio corrispondente all'autovalore λ_n è il sottospazio generato da x_n. Per quanto detto a pag. 309 si ha quindi:

$$P_{\lambda_m} - P_{\lambda_{m+1}} = P_n \quad . \tag{9}$$

3) Risulta $m=0$. Infatti essendo $m \leq \lambda_n \ \forall n \in N$ [(3.1) pag. 242] e $\lim \lambda_n = 0$, si ha $m \leq 0$: di conseguenza m non è autovalore di A (gli autovalori sono positivi). Allora m, punto dello spettro per la citata (3.1), appartiene allo spettro continuo di A [lo spettro residuo è vuoto, ancora per la (3.1)]: per la (3.2) di pag. 244 ne segue che $m=0$.

Ciò premesso, ricordiamo che se $f(\lambda)$ è una funzione continua in un intervallo $[a,b]$, e se $\varphi(\lambda)$ è una funzione costante in $[a,b[$ definita anche in b, si ha [

$$\int_a^b f(\lambda)\, d\varphi = f(b)\,[\varphi(b) - \varphi(a)] \, .$$

Fissati $x, y \in S$ e posto $\varphi(\lambda) = (P_\lambda x, y) \ \forall \lambda \in R$, essendo $\varphi(\lambda)$ costante in $[\lambda_{m+1}, \lambda_n[$ si ha allora, tenendo presente la (9):

$$\int_{\lambda_{m+1}}^{\lambda_m} \lambda \, d\varphi = \lambda_n [\varphi(\lambda_m) - \varphi(\lambda_{m+1}) = \lambda_n \left[(P_{\lambda_m} x, y) - (P_{\lambda_{m+1}} x, y) \right] = \lambda_n \left((P_{\lambda_m} - P_{\lambda_{m+1}}) x, y \right) = \lambda_n (P_n x, y)$$

e quindi:

$$\int_{\lambda_{m+1}}^{\lambda_m} \lambda \, d\varphi = (\lambda_n P_n x, y) \, . \tag{10}$$

D'altra parte, per la (4) di pag. 307 si ha:

$$\int_{-\infty}^{+\infty} \lambda \, d(P_\lambda x, y) = \int_{-\infty}^{+\infty} \lambda \, d\varphi = m\varphi(m) + \int_m^M \lambda \, d\varphi = \int_0^M \lambda \, d\varphi = \sum_{n=1}^{\infty} \int_{\lambda_{n+1}}^{\lambda_n} \lambda \, d\varphi$$

e quindi per la (10):

$$\int_{-\infty}^{+\infty} \lambda \, d(P_\lambda x, y) = \left(\sum_{n=1}^{\infty} \lambda_n P_n x, y \right) \tag{11}$$

Orbene, la (11) implica l'uguaglianza:

$$\int_{-\infty}^{+\infty} \lambda \, d P_\lambda = \sum_{n=1}^{\infty} \lambda_n P_n \tag{12}$$

cioè l'assento. Infatti il primo membro di (12) rappresenta [pag. 307] l'operatore che ad ogni $x \in S$ associa l'elemento $\tilde{x}$ di S tale che

$$(\tilde{x}, y) = \int_{-\infty}^{+\infty} \lambda \, d(P_\lambda x, y) \qquad \forall y \in S$$

ossia, per la (11), tale che:

$$(\tilde{x}, y) = \left(\sum_{n=1}^{\infty} \lambda_n P_n x, y \right) \qquad \forall y \in S \, .$$

Siccome da tale relazione consegue $\tilde{x} = \sum_{n=1}^{\infty} \lambda_n P_n x$, si conclude che ad ogni $x \in S$ l'operatore a primo membro di (11) associa $\sum_{n=1}^{\infty} \lambda_n P_n x$, sicchè l'operatore coincide con $\sum_{n=1}^{\infty} \lambda_n P_n$.

Supposto ancora che l'operatore A sia autoaggiunto, torniamo a considerare l'operatore (6), lineare e continuo:

$$B = \int_{-\infty}^{+\infty} f(\lambda)\, dP_\lambda ,$$

scegliendo ivi come misura spettrale $(P_\lambda)_{\lambda\in R}$ quella associata all'operatore A. I numeri m, M relativi a $(P_\lambda)_{\lambda\in R}$ sono allora gli estremi dell'operatore A.

Evidentemente l'operatore B dipende da A, per il tramite della misura spettrale $(P_\lambda)_{\lambda\in R}$; d'altronde e' proprio uguale ad A se $f(\lambda)=\lambda$. Naturalmente B dipende anche dalla scelta della funzione $f(\lambda)$, a valori reali o a valori complessi, definita e continua nell'intervallo $[m,M]$.

Il legame tra l'operatore B e l'operatore A e' molto espressivo dal punto di vista formale se si sceglie $f(\lambda)=\lambda^n$ oppure $f(\lambda)=(\lambda-\lambda_0)^{-1}$ con $\lambda_0\in\rho(A)$, e se si conviene di denotare l'operatore $A-\lambda I$ semplicemente con $A-\lambda$. Infatti si dimostra che:

$$\text{se } f(\lambda)=\lambda^n \text{ si ha } B=A^n$$
$$\text{se } f(\lambda)=(\lambda-\lambda_0)^{-1} \text{ si ha } B=(A-\lambda_0)^{-1}$$

ed e' ovvio che per ogni $\kappa\in C$:

$$\text{se } f(\lambda)=\kappa\lambda \text{ si ha } B=\kappa A .$$

Dunque, nei casi suddetti, l'operatore B si ottiene dall'espressione di $f(\lambda)$ sostituendo formalmente λ con A. Ciò suggerisce di denotare in ogni caso l'operatore B con $f(A)$, cioe' di porre:

$$f(A) = \int_{-\infty}^{+\infty} f(\lambda)\, dP_\lambda , \tag{9}$$

e quindi ad esempio:

$$\cos A = \int_{-\infty}^{+\infty} \cos\lambda\, dP_\lambda \quad , \quad \operatorname{sen} A = \int_{-\infty}^{+\infty} \operatorname{sen}\lambda\, dP_\lambda \quad , \quad e^{iA} = \int_{-\infty}^{+\infty} e^{i\lambda}\, dP_\lambda \quad , \text{ ecc.}$$

In virtu' della proposiz. II, l'operatore (9) e' autoaggiunto se e solo se $f(\lambda)$ e' reale. Notiamo che l'operatore

$$\sqrt{A} = \int_{-\infty}^{+\infty} \sqrt{\lambda}\, dP_\lambda \qquad \left[\text{risp. } \log A = \int_{-\infty}^{+\infty} \log\lambda\, dP_\lambda\right]$$

e' autoaggiunto se $\lambda\geq 0$ [risp. $\lambda>0$] nell'intervallo $[m,M]$, cioe' se l'operatore A (autoaggiunto per ipotesi) e' positivo [risp. strettamente positivo].

Si dimostra che, qualunque siano f, g definite e continue in $[m,M]$, risulta:

$$f\cdot g\,(A) = f(A)\,g(A) ,$$

ossia:

$$f(A)\,g(A) = \int_{-\infty}^{+\infty} f(\lambda)\,g(\lambda)\, dP_\lambda .$$

In particolare si ha:

$$\sqrt{A}\cdot\sqrt{A} = A .$$

Si dimostra inoltre che, se f è reale e risulta $f(\lambda)\geq 0 \ \forall \lambda \in [m,M]$, allora $f(A)\geq 0$.

In particolare si ha, ad esempio.

$$e^A > 0 \ .$$

Riportiamo infine la relazione:

$$\|f(A)\| \leq \max_{\lambda\in[m,M]} |f(\lambda)| \ .$$

Consideriamo ora un operatore unitario U di S in S, ossia un operatore, lineare e continuo, tale che:

$$UU^* = U^*U = I \ ,$$

ossia tale che $U^* = U^{-1}$. Come sappiamo, tale proprietà caratterizza gli isomorfismi isometrici di S in S. Orbene, sussiste il teorema:

IV. Per ogni operatore unitario U di S in S, esiste una e una sola misura spettrale $(P_\lambda)_{\lambda\in R}$ tale che:

$$U = \int_{-\infty}^{+\infty} e^{2\pi i\lambda}\, dP_\lambda \ . \tag{10}$$

La misura spettrale $(P_\lambda)_{\lambda\in R}$ per la quale sussiste la (10) determina un operatore autoaggiunto, diciamolo B, cioè poniamo:

$$B = \int_{-\infty}^{+\infty} \lambda\, dP_\lambda \ .$$

Dalla (10) consegue allora:

$$e^{2\pi i B} = U$$

Pertanto la proposiz. IV si può enunciare equivalentemente come segue: per ogni operatore unitario U di S in S, esiste uno e un solo operatore autoaggiunto B tale che:

$$U = e^{2\pi i B} \ .$$

Concludiamo rilevando che la posizione (9) si può estendere a funzioni non necessariamente continue: alla (9) si può dare significato anche se $f(\lambda)$ è limitata nonché misurabile rispetto ad una opportuna misura μ legata alla famiglia di proiettori $(P_\lambda)_{\lambda\in R}$. Non possiamo qui soffermarci sulla questione. Ci limitiamo a dire che, dopo la generalizzazione accennata, se la funzione $f(\lambda)$ assume soltanto i valori 0 e 1 l'operatore $f(A)$ dato dalla (9) è un proiettore.

Perturbazioni delle disequazioni variazionali

Siano: X uno spazio di Banach riflessivo, K un sottoinsieme chiuso e convesso (non vuoto) di X.

Dato $A: X \to X'$ un operatore monotono ed unicontinuo, sia $\mathcal{K}$ l'insieme delle soluzioni della disequazione variazionale:

(1) $\qquad u \in K: \ \langle Au, v-u \rangle \geq 0 \qquad \forall v \in K.$

Non avendo supposto K limitato, e neppure che l'operatore A sia coercivo, non si può dire che sono soddisfatte le ipotesi del teor. (1.6) di pag. 257.3

Fissiamo un operatore $B: X \to X'$ strettamente monotono, unicontinuo e tale che:

(2) $\qquad \langle Bu - Bv, u-v \rangle \geq \|u-v\| \, \gamma(\|u-v\|) \qquad \forall u, v \in X,$

essendo $\gamma(\tau)$ una funzione reale, continua e crescente in $[0, +\infty[$, positiva per $\tau > 0$, nulla per $\tau = 0$ e divergente per $\tau \to +\infty$. In tali ipotesi su B, quando risulta $\mathcal{K} \neq \emptyset$ la disequazione variazionale:

(3) $\qquad u \in \mathcal{K}: \ \langle Bu, v-u \rangle \geq 0 \quad \forall v \in \mathcal{K}$

ammette una e una sola soluzione.

Per ogni $\varepsilon > 0$ l'operatore:

$$A_\varepsilon = A + \varepsilon B$$

si chiama una <u>perturbazione</u> di A, e la disequazione variazionale:

(4) $\qquad u \in K: \ \langle A_\varepsilon u, v-u \rangle \geq 0 \quad \forall v \in K$

si chiama una <u>perturbazione</u> della (1). Evidentemente l'operatore A_ε soddisfa alle stesse ipotesi poste su B, e quindi la (4) ammette una e una sola soluzione, che denoteremo con u_ε.

Orbene, anche se K non è limitato, ma è limitato l'insieme (contenuto in K) degli elementi della famiglia $(u_\varepsilon)_{\varepsilon \in]0, +\infty[}$, la disequazione variazionale (1) ammette almeno una soluzione; sussiste invero il teorema:

I. La disequazione variazionale (1) ammette almeno una soluzione se e solo se esiste una costante positiva L tale che:

(5) $\qquad \|u_\varepsilon\| \le L \quad \forall \varepsilon > 0 .$

Quando tale condizione e' soddisfatta (in particolare se K e' limitato), u_ε converge per $\varepsilon \to 0$ ad una soluzione della (1): precisamente tende all'unica soluzione u^* della (3).

DIM. Supposto che la (1) ammetta almeno una soluzione, ossia che $K \neq \emptyset$, denotiamo con u_0 un qualunque elemento di K.

Risulta allora:

(6) $\qquad \langle Au_0, v-u_0 \rangle \geq 0 \quad \forall v \in K ,$

e d'altra parte, essendo u_ε la soluzione della (4), si ha:

(7) $\qquad \langle Au_\varepsilon + \varepsilon Bu_\varepsilon, v-u_\varepsilon \rangle \geq 0 \quad \forall v \in K .$

Ponendo in (6) $v = u_\varepsilon$ e in (7) $v = u_0$, indi sommando membro a membro, si ottiene:

$$-\langle Au_\varepsilon - Au_0, u_\varepsilon - u_0 \rangle + \varepsilon \langle Bu_\varepsilon, u_0 - u_\varepsilon \rangle \geq 0$$

da cui, essendo ≤ 0 il primo addendo al primo membro, segue:

(8) $\qquad \langle Bu_\varepsilon, u_0 - u_\varepsilon \rangle \geq 0 .$

Tenendo conto della (8), dalla (2) si deduce:

$$\gamma(\|u_\varepsilon - u_0\|)\|u_\varepsilon - u_0\| \le \langle Bu_\varepsilon - Bu_0, u_\varepsilon - u_0 \rangle \le -\langle Bu_0, u_\varepsilon - u_0 \rangle \le \|Bu_0\| \cdot \|u_\varepsilon - u_0\|$$

e quindi risulta:

$$\gamma(\|u_\varepsilon - u_0\|) \le \|Bu_0\| .$$

Poiché il secondo membro e' indipendente da ε, per le ipotesi su γ ne segue che esiste $h > 0$ tale che $\|u_\varepsilon - u_0\| \le h$, e quindi e' vera la (5) con $L = \|u_0\| + h$.

Viceversa, supponiamo vera la (5). Esiste allora una successione infinitesima $\{\varepsilon_n\}$ di numeri positivi tale che $\{u_{\varepsilon_n}\}$ converga debolmente in X ad un elemento $w \in K$ (1).

(1) Essendo K un insieme convesso, l'ipotesi K chiuso equivale al fatto che K e' debolmente chiuso. Pertanto, appartenendo w alla chiusura debole di K (in quanto limite debole di una successione di elementi di K), si ha $w \in K$.

Essendo u_{ε_n} soluzione della (4) per $\varepsilon=\varepsilon_n$, applicando alla (4) il lemma di Minty si ha:

$$\langle A_{\varepsilon_n} v, v-u_{\varepsilon_n}\rangle \geq 0 \quad \forall v\in K,$$

da cui, passando al limite per $n\to\infty$, segue:

$$\langle Av, v-w\rangle \geq 0 \quad \forall v\in K.$$

Applicando ancora il lemma di Minty, se ne deduce che w è soluzione della (1), e con ciò la prima parte del teorema è dimostrata.

Per dimostrare la seconda parte, supponiamo vera la (5) e consideriamo la soluzione u^* della (3), ricché $u^*\in K$ e risulta:

$$\langle Bu^*, v-u^*\rangle \geq 0 \quad \forall v\in K. \tag{9}$$

Detta $\{\varepsilon_n\}$ una qualunque successione infinitesima di numeri positivi tale che $\{u_{\varepsilon_n}\}$ sia debolmente convergente, dal ragionamento prima svolto si deduce che il limite debole w di $\{u_{\varepsilon_n}\}$ è una soluzione della (1), ossia è un elemento di K, ed inoltre che [cfr. (8)]:

$$\langle Bu_\varepsilon, u_\varepsilon-u_0\rangle \leq 0 \quad \forall u_0\in K. \tag{10}$$

Cominciamo col dimostrare che $w=u^*$. A tale scopo osserviamo che, posto $\zeta=(1-t)w+tu^*$, ossia:

$$w-\zeta = t(w-u^*),$$

tenendo presente che B è monotono e utilizzando la (10) prima per $u_0=w$, poi per $u_0=u^*$, si ha:

$$\langle B\zeta, u_{\varepsilon_n}-\zeta\rangle \leq \langle Bu_{\varepsilon_n}, u_{\varepsilon_n}-\zeta\rangle = \langle Bu_{\varepsilon_n}, u_{\varepsilon_n}-w\rangle + \langle Bu_{\varepsilon_n}, w-\zeta\rangle$$
$$\leq \langle Bu_{\varepsilon_n}, w-\zeta\rangle = t\langle Bu_{\varepsilon_n}, w-u^*\rangle = t\left[\langle Bu_{\varepsilon_n}, w-u_{\varepsilon_n}\rangle + \langle Bu_{\varepsilon_n}, u_{\varepsilon_n}-u^*\right]$$
$$\leq t\langle Bu_{\varepsilon_n}, w-u_{\varepsilon_n}\rangle \leq t\langle Bw, w-u_{\varepsilon_n}\rangle$$

da cui, passando al limite per $n\to\infty$, si trae $\langle B\zeta, w-\zeta\rangle \leq 0$, ossia:

$$\langle B\zeta, w-u^*\rangle \leq 0.$$

Essendo B emicontinuo, passando al limite per $t\to 0$ si ottiene:

$$\langle Bw, w-u^*\rangle \leq 0. \tag{11}$$

e d'altra parte la (9), per $v=w$, fornisce:

$$\langle -Bu^*, w-u^*\rangle \leq 0. \tag{12}$$

Sommando membro a membro (11) e (12) si ha:

$$\langle Bw-Bu^*, w-u^*\rangle \leq 0,$$

da cui, essendo B strettamente monotono, si trae che $w = u^*$.

Osserviamo ora che, tenendo conto della (10) con $u_0 = u^*$, dalla (2) si ha:

$$0 \le \gamma(\|u_{\varepsilon_n} - u^*\|)\,\|u_{\varepsilon_n} - u^*\| \le \langle Bu_{\varepsilon_n} - Bu^*, u_{\varepsilon_n} - u^* \rangle \le \langle -Bu^*, u_{\varepsilon_n} - u^* \rangle$$

e quindi, essendo u^* il limite debole di $\{u_{\varepsilon_n}\}$:

$$\lim_{n\to\infty} \gamma(\|u_{\varepsilon_n} - u^*\|)\,\|u_{\varepsilon_n} - u^*\| = 0 .$$

Per le proprietà della funzione γ ne segue che:

$$\lim_{n\to\infty} \|u_{\varepsilon_n} - u^*\| = 0 .$$

Abbiamo dunque stabilito che, qualunque sia la successione infinitesima $\{\varepsilon_n\}$ di numeri positivi, dalla successione $\{u_{\varepsilon_n}\}$ se ne può estrarre una convergente (fortemente) ad u^*. Ciò equivale al fatto che:

$$\lim_{\varepsilon\to 0} u_\varepsilon = u^*$$

e il teorema è dimostrato.

* * *

Supponiamo in particolare che X sia uno spazio di Hilbert, e denotiamo ancora con K un sottoinsieme chiuso e convesso di X.

Sia $a(u,v)$ una forma bilineare su X, continua e tale che $a(v,v) \ge 0 \ \forall v \in X$. Detto f' un elemento di X', sia $\mathcal{K}$ l'insieme delle soluzioni della disequazione variazionale:

(13) $\qquad u \in K : a(u, v-u) \ge \langle f, v-u \rangle \quad \forall v \in K .$

Non avendo supposto $a(u,v)$ coercitiva, l'insieme $\mathcal{K}$ può essere vuoto, ma è ovviamente chiuso e convesso.

Fissati un elemento g di X' ed una forma bilineare $b(u,v)$ su X, continua e coercitiva, per ogni $\varepsilon > 0$ la disequazione variazionale:

(14) $\qquad u \in K : a(u, v-u) + \varepsilon\, b(u, v-u) \ge \langle f + \varepsilon g, v-u \rangle \quad \forall v \in K$

ammette una e una sola soluzione u_ε (1).

(1) Essendo $a(v,v) + \varepsilon\, b(v,v) \ge \varepsilon\, b(v,v)$, e $b(u,v)$ coercitiva, la forma bilineare $a(u,v) + \varepsilon\, b(u,v)$ è coercitiva (oltre che continua).

Denotando con A e B gli operatori di X in X' rispettivamente definiti da:

$$\langle Au, v\rangle = a(u,v) - \langle f, v\rangle \quad , \quad \langle Bu, v\rangle = b(u,v) - \langle g, v\rangle \quad ,$$

tali operatori soddisfano alle ipotesi poste per il teorema I. Si ha allora, in particolare, il seguente teorema:

II. La disequazione variazionale (13) ammette almeno una soluzione se e solo se esiste una costante positiva L tale che:

$$\|u_\varepsilon\| \leq L \qquad \forall \varepsilon > 0 .$$

Quando tale condizione è soddisfatta (in particolare se K è limitato), u_ε converge per $\varepsilon \to 0$ ad una soluzione della (13): precisamente tende all'unica soluzione u^* della disequazione variazionale:

$$(15) \qquad u \in K : \; b(u, v-u) \geq \langle g, v-u\rangle \quad \forall v \in K .$$

Un teorema astratto di regolarità per le soluzioni delle disequazioni variazionali.

Siano: V uno spazio di Banach riflessivo, V' il suo duale, K un sottoinsieme chiuso e convesso di V, A un operatore di V in V', monotono ed emicontinuo, f un elemento di V'. Se K non è limitato, supporremo che A soddisfi alla condizione:

$$(1) \qquad \frac{\langle Av, v\rangle}{\|v\|} \to +\infty \quad \text{per } \|v\| \to +\infty .$$

Nelle ipotesi poste la disequazione variazionale:

$$(2) \qquad u \in K : \ \langle Au - f, v - u\rangle \geq 0 \quad \forall v \in K$$

ammette almeno una soluzione: indichiamo con u^* una qualunque delle soluzioni della (2).

Consideriamo un sottospazio vettoriale W di V' denso in V', che si suppone normato e con immersione continua in V'.

Se u^* è interno all'insieme K, la (2) con $u = u^*$ si riduce all'uguaglianza $\langle Au^* - f, v\rangle = 0 \ \forall v \in V$, e quindi implica che $Au^* = f$. Pertanto è banale che:

$$(3) \qquad f \in W \Longrightarrow Au^* \in W .$$

Se invece u^* non è interno a K, non si può asserire che Au^* coincida con f, e quindi non è detto che sussista la (3).

Orbene, vogliamo stabilire delle condizioni atte ad assicurare la validità della (3): quest'ultima comporta in sostanza che, se l'elemento f di V' gode di particolari proprietà (dovute all'appartenenza allo spazio W), anche Au^* gode delle stesse proprietà, e da ciò si può risalire (se si conoscono proprietà delle soluzioni dell'equazione $Au = \eta$ in dipendenza di proprietà del termine noto η) a proprietà di u^* (in quanto, posto $Au^* = \eta$, u^* è soluzione dell'equazione $Au = \eta$). In altri termini, tramite la (3) si possono ottenere teoremi "di regolarità" per le soluzioni della (2), ossia teoremi che in particolari ipotesi su f assicurano particolari proprietà di u^*.

Supponiamo che W, e quindi anche il suo duale W', sia uno spazio di Banach riflessivo. Supporremo inoltre, ogni qualvolta non vi sia possibilità di equivoci, di aver identificato lo spazio V col suo biduale V'', sicché, essendo $V'' \subseteq W'$, lo spazio V sarà considerato come sottospazio di W'.

Indicheremo con $\langle \cdot, \cdot \rangle$ la dualità tra V' e V, e con $\langle\langle \cdot, \cdot \rangle\rangle$ la dualità tra W' e W.

Fissata una funzione reale $\varphi(r)$, definita in $[0, +\infty[$, strettamente crescente, nulla per $r=0$ e divergente per $r \to +\infty$, denoteremo con J l'applicazione di dualità di W in W' determinata da φ (1).

* * *

Per ogni $\varepsilon > 0$ consideriamo due applicazioni:

(4) $\qquad B_\varepsilon : V \to W \quad , \quad B'_\varepsilon : V \to W'$

tali che si abbia, per un'opportuna costante b indipendente da ε:

(5) $\qquad \|B_\varepsilon(v)\|_W \leq b \quad , \quad \|B'_\varepsilon(v)\|_{W'} \leq b \qquad \forall v \in V.$

Supponiamo verificata la seguente condizione:

(6) <u>Per ogni $\varepsilon > 0$ esiste un unico $u_\varepsilon \in K$ tale che:</u>

(6_1) $\qquad A u_\varepsilon \in W$

(6_2) $\qquad u_\varepsilon + \varepsilon J\left[(A + B_\varepsilon) u_\varepsilon + B'_\varepsilon u_\varepsilon\right] = u^*$ (2).

Le ipotesi poste individuano, in corrispondenza di u^*, una famiglia $\{u_\varepsilon\}$ di elementi di K che gode di particolari proprietà.

---

(1) Si dimostra che la funzione φ individua una funzione $J : W \to W'$, detta <u>applicazione di dualità</u>, tale da risultare per ogni $w \in W$:

$$\langle\langle Jw, w \rangle\rangle = \varphi(\|w\|)\|w\| \quad , \quad \|Jw\| = \varphi(\|w\|).$$

(2) Si noti che per la (6_1) il coefficiente di ε in (6_2) è un elemento di W' e che u_ε, u^* sono elementi di $V \subseteq W'$. Dell'esistenza di un $u_\varepsilon \in V$ soddisfacente alle (6_1), (6_2) ci occuperemo in seguito.

Esprimiamo tali proprietà mediante i teoremi seguenti, dei quali il terzo è conseguenza del secondo, e a sua volta il secondo è conseguenza del primo.

I. Se $f \in W'$, esiste una costante c tale che:

(7) $$\|Au_\varepsilon\|_{W'} \leq c \qquad \forall \varepsilon > 0$$

e risulta:

(8) $$\lim_{\varepsilon \to 0} u_\varepsilon = u^* \quad \text{fortemente in } W'.$$

II. Se $f \in W'$, si ha:

(9) $$\lim_{\varepsilon \to 0} \langle Au_\varepsilon, u_\varepsilon - u^* \rangle = 0,$$

ed esiste una sottofamiglia $\{u_{\varepsilon'}\}$ di $\{u_\varepsilon\}$ tale che:

(10) $$\lim_{\varepsilon' \to 0} u_{\varepsilon'} = u^* \quad \text{debolmente in } V.$$

III. Se $f \in W'$, esiste una sottofamiglia $\{u_{\varepsilon'}\}$ di $\{u_\varepsilon\}$ tale che:

$$\lim_{\varepsilon' \to 0} Au_{\varepsilon'} = Au^* \quad \text{debolmente in } V'.$$

DIM. I.

Essendo u^* soluzione della (1), per il lemma di Minty [1] si ha $\langle Av - f, v - u^* \rangle \geq 0 \quad \forall v \in K$, e quindi, per $v = u_\varepsilon$:

(11) $$\langle Au_\varepsilon - f, u_\varepsilon - u^* \rangle \geq 0.$$

Posto:

(12) $$z_\varepsilon = (A + B_\varepsilon) u_\varepsilon,$$

risulta:

$$Au_\varepsilon - f = z_\varepsilon - B_\varepsilon u_\varepsilon - f,$$

e d'altra parte la (6_2) può scriversi:

(13) $$u_\varepsilon - u^* = -\varepsilon \left[J(z_\varepsilon) + B'_\varepsilon u_\varepsilon \right].$$

Dalla (11) consegue allora:

(14) $$\langle z_\varepsilon - B_\varepsilon u_\varepsilon - f, J(z_\varepsilon) + B'_\varepsilon u_\varepsilon \rangle \leq 0.$$

Osserviamo che per la (13) la somma $J(z_\varepsilon) + B'_\varepsilon u_\varepsilon$ è un elemento

---

(1) Cfr. (1.8) pag. 257.4

di V, e quindi anche di W'; per quanto attiene ai singoli addendi, non si può dire che appartengono a V, ma sono elementi di $W' \supseteq V$. Poichè l'elemento $z_\varepsilon - B_\varepsilon u_\varepsilon - f$ di V' appartiene più precisamente a $W \subseteq V'$, la (14) può scriversi:

$$0 \geq \langle\langle J(z_\varepsilon) + B'_\varepsilon u_\varepsilon \,,\, z_\varepsilon - B_\varepsilon u_\varepsilon - f \rangle\rangle = \langle\langle J(z_\varepsilon), z_\varepsilon - B_\varepsilon u_\varepsilon - f \rangle\rangle + \langle\langle B'_\varepsilon u_\varepsilon, z_\varepsilon - B_\varepsilon u_\varepsilon - f \rangle\rangle$$

cioè:

$$\langle\langle J(z_\varepsilon), z_\varepsilon \rangle\rangle - \langle\langle J(z_\varepsilon), B_\varepsilon u_\varepsilon + f \rangle\rangle + \langle\langle B'_\varepsilon u_\varepsilon \,,\, z_\varepsilon - B_\varepsilon u_\varepsilon - f \rangle\rangle \leq 0 \,,$$

ossia:

$$\langle\langle J(z_\varepsilon), z_\varepsilon \rangle\rangle \leq \langle\langle J(z_\varepsilon), B_\varepsilon u_\varepsilon + f \rangle\rangle + \langle\langle -B'_\varepsilon u_\varepsilon, z_\varepsilon - B_\varepsilon u_\varepsilon - f \rangle\rangle \,.$$

Pertanto, ricordando quanto detto in (*) a pag. 318 e tenendo presente la (5), si ha:

$$\varphi(\|z_\varepsilon\|)\,\|z_\varepsilon\| \leq \varphi(\|z_\varepsilon\|)\,(b + \|f\|) + b\,(\|z_\varepsilon\| + b + \|f\|) \,,$$

le norme essendo tutte in W. Posto $b_1 = b + \|f\|$, ne segue che:

$$\varphi(\|z_\varepsilon\|)\,(\|z_\varepsilon\| - b_1) \leq b\,(\|z_\varepsilon\| + b_1) \,. \tag{15}$$

Dalla (15) si deduce l'esistenza di una costante k tale che:

$$\|z_\varepsilon\| \leq k \quad \forall \varepsilon > 0 \,. \tag{16}$$

Ciò è banale se $\|z_\varepsilon\| \leq b_1 \ \forall \varepsilon > 0$, e quindi supponiamo che esista almeno un $\varepsilon > 0$ tale che $\|z_\varepsilon\| > b_1$. Per ogni ε siffatto la (15) si può scrivere:

$$\varphi(\|z_\varepsilon\|) \leq b\,\frac{\|z_\varepsilon\| + b_1}{\|z_\varepsilon\| - b_1} \,. \tag{17}$$

Procedendo per assurdo, ammettiamo che non esista alcun $k > 0$ tale da verificare la (16): allora $\forall n \in \mathbb{N}$ esiste $\varepsilon_n > 0$ tale che $\|z_{\varepsilon_n}\| \geq n + b_1$; posto $\|z_{\varepsilon_n}\| = z_n$ si ha dunque $\lim_n z_n = +\infty$, nonché per la (17):

$$\varphi(z_n) \leq b\,\frac{z_n + b_1}{z_n - b_1} \,.$$

Ne segue:

$$\lim_n \varphi(z_n) \leq b \,,$$

il che è assurdo essendo $\varphi(z)$ divergente per $z \to +\infty$.

Dalla (16) ora acquisita, ricordando le (12) e (5), si trae:

$$\|Au_\varepsilon\| \leq \|Au_\varepsilon + B_\varepsilon u_\varepsilon\| + \|B_\varepsilon u_\varepsilon\| = \|z_\varepsilon\| + \|B_\varepsilon u_\varepsilon\| \leq k + b \,,$$

e la (7) è dimostrata.

Per dimostrare la (8) osserviamo che, essendo $\|Jw\| = \varphi(\|w\|)$ per ogni $w \in W$, l'applicazione J trasforma insiemi limitati di W in insiemi limitati di W', e quindi per le (5) e (7) esiste una costante h tale che:

$$\| J(A+B_\varepsilon)u_\varepsilon + B'_\varepsilon u_\varepsilon \|_{W'} \leq h \quad , \forall \varepsilon > 0 .$$

Dopo ciò la (8) è ovvia, in quanto per la (6_2) si ha:

(18) $$\|u^* - u_\varepsilon\|_{W'} \leq h\varepsilon .$$

DIM. II.

La (9) si ricava subito dal teorema I. Basta osservare che:

$$|\langle Au_\varepsilon, u_\varepsilon - u^* \rangle| = |\langle\langle u_\varepsilon - u^*, Au_\varepsilon \rangle\rangle| \leq \|u_\varepsilon - u^*\|_{W'} \cdot \|Au_\varepsilon\|_W \leq c \|u_\varepsilon - u^*\|_{W'} .$$

Per dimostrare la seconda parte del teorema, proviamo intanto che esiste $\delta > 0$ tale che la famiglia $\{u_\varepsilon\}_{\varepsilon \in]0,\delta[}$ è limitata in V. Ciò è banale se K è limitato (in quanto $u_\varepsilon \in K \ \forall \varepsilon > 0$); supposto K non limitato, nel qual caso l'operatore A soddisfa alla (1), osserviamo che:

$$\langle Au_\varepsilon, u_\varepsilon \rangle = \langle Au_\varepsilon, u_\varepsilon - u^* \rangle + \langle Au_\varepsilon, u^* \rangle ,$$

e che dei due addendi al secondo membro il primo tende a zero per $\varepsilon \to 0$ (per la (9)), mentre il secondo è limitato (per la (7), ricordando che l'immersione di W in V' è continua). Ne segue che esiste un $\delta > 0$ tale che per $\varepsilon \in]0,\delta[$ l'insieme numerico descritto dal primo membro è limitato. Pertanto si ha:

(19) $$|\langle Au_\varepsilon, u_\varepsilon \rangle| \leq a_0 \quad \forall \varepsilon \in]0,\delta[,$$

con a_0 costante indipendente da ε.

Dalla (19) si deduce l'esistenza di una costante K tale che:

$$\|u_\varepsilon\| \leq K \quad \forall \varepsilon \in]0,\delta[,$$

in quanto, se così non fosse, per la (19) verrebbe contraddetta l'ipotesi (1).

Dunque esiste una sottofamiglia $\{u_{\varepsilon'}\}$ di $\{u_\varepsilon\}$ debolmente convergente in V, per $\varepsilon' \to 0$, verso un elemento $u_0 \in V$:

$$\lim_{\varepsilon' \to 0} \langle z, u_{\varepsilon'} \rangle = \langle z, u_0 \rangle \quad \forall z \in V' \text{ (e quindi anche } \forall z \in W \subseteq V') .$$

Ne segue:

$$\lim_{\varepsilon'\to 0} \langle\langle u_{\varepsilon'}, w\rangle\rangle = \langle\langle u_0, w\rangle\rangle \qquad \forall w \in W . \tag{20}$$

D'altra parte, per la (8) si ha:

$$\lim_{\varepsilon'\to 0} \langle\langle u_{\varepsilon'}, w\rangle\rangle = \langle\langle u^*, w\rangle\rangle \qquad \forall w \in W \ (1) . \tag{21}$$

Dalle (20) e (21) consegue:

$$\langle\langle u_0, w\rangle\rangle = \langle\langle u^*, w\rangle\rangle$$

per ogni $w \in W$, e quindi anche per ogni $w \in V'$ (essendo W denso in V'). Pertanto $u_0 = u^*$ e con ciò il teorema II è dimostrato.

DIM. III. (2).

Detto v un elemento di V, e posto:

$$w = u^* + t(v - u^*) \qquad \forall t \in]0,1] ,$$

essendo A monotono risulta:

$$0 \le \langle Au_{\varepsilon'} - Aw, u_{\varepsilon'} - w\rangle =$$

$$= \langle Au_{\varepsilon'}, u_{\varepsilon'} - u^*\rangle - t\langle Au_{\varepsilon'}, v - u^*\rangle - \langle Aw, u_{\varepsilon'} - w\rangle ,$$

ossia:

$$t\langle Au_{\varepsilon'}, v - u^*\rangle \le \langle Au_{\varepsilon'}, u_{\varepsilon'} - u^*\rangle + \langle Aw, w - u_{\varepsilon'}\rangle .$$

Poiché per $\varepsilon' \to 0$ si ha $u_{\varepsilon'} \to u^*$ debolmente in V [cfr. (10)], e il primo addendo al secondo membro tende a 0 [cfr. (9)], ne segue:

$$t \lim''_{\varepsilon'\to 0} \langle Au_{\varepsilon'}, v - u^*\rangle \le \langle Aw, w - u^*\rangle = t\langle Aw, v - u^*\rangle$$

---

(1) In quanto $|\langle\langle u_{\varepsilon'} - u^*, w\rangle\rangle| \le \|u_{\varepsilon'} - u^*\|_{W'} \cdot \|w\|_W$.

(2) E' utile osservare che, come risulterà dalla dimostrazione, il teorema III sussiste per una qualunque famiglia $\{u_{\varepsilon'}\}$ di elementi di V tale che:

$$\lim_{\varepsilon'\to 0} u_{\varepsilon'} = u^* \text{ debolmente in } V , \quad \lim_{\varepsilon'\to 0} \langle Au_{\varepsilon'}, u_{\varepsilon'} - u^*\rangle = 0 ,$$

qualunque sia l'operatore $A: V \to V'$ monotono ed emicontinuo [ossia, indipendentemente dal fatto che $u_{\varepsilon'} \in K$ e soddisfa alle $(6_1), (6_2)$]. Nel teorema III l'ipotesi $f \in W$ serve solo ad assicurare le soprascritte relazioni [cfr. teor. II].

e quindi, dividendo per t e passando al limite per $t \to 0$:

$$(22) \qquad \lim_{\varepsilon' \to 0}{}'' \langle Au_{\varepsilon'}, v-u^* \rangle \leq \langle Au^*, v-u^* \rangle .$$

La (22) sussiste qualunque sia $v \in V$, e quindi anche se in essa si sostituisce v con $2u^*-v$, per cui la (22) conduce alla seguente:

$$(23) \qquad \langle Au^*, v-u^* \rangle \leq \lim_{\varepsilon' \to 0} \langle Au_{\varepsilon'}, v-u^* \rangle .$$

Da (22) e (23) segue:

$$\lim_{\varepsilon' \to 0} \langle Au_{\varepsilon'}, v-u^* \rangle = \langle Au^*, v-u^* \rangle$$

e quindi ponendo, qualunque sia $z \in V$, $v = u^* + z$:

$$\lim_{\varepsilon' \to 0} \langle Au_{\varepsilon'}, z \rangle = \langle Au^*, z \rangle \qquad \forall z \in V .$$

Con ciò il teorema è dimostrato.

* * *

Vogliamo ora mostrare che è possibile scegliere gli operatori B_ε, B'_ε in modo che esista un unico elemento $u_\varepsilon \in V$ tale da soddisfare alle (6_1), (6_2), almeno nell'ipotesi che A soddisfi alla (1) (anche nel caso K limitato) e che gli spazi W, W' siano strettamente convessi (*).

Il fatto poi che u_ε appartenga al convesso K dipende, naturalmente, dal particolare convesso che si considera.

Sia B_ε un operatore di V in V' tale che $B_\varepsilon(V) \subseteq W$, monotono ed emicontinuo; detta b_ε una costante positiva, definiamo l'operatore B'_ε come segue:

$$B'_\varepsilon : v \in V \to b_\varepsilon v \in V \subseteq W' ,$$

e supponiamo verificata la (5) [con b costante indipendente da ε].

Consideriamo poi l'applicazione inversa di J, la cui restrizione a V è a valori in V' (in quanto $W \subseteq V'$), e sarà denotata con J^{-1}. Essendo W e W' strettamente convessi, l'operatore J^{-1} è strettamente monotono, ed inoltre tale operatore è emicontinuo.

Ciò premesso, per ogni $\varepsilon > 0$ poniamo:

(*) Uno spazio vettoriale normato S si dice <u>strettamente convesso</u> se per ogni coppia (x, y) di punti distinti di S, aventi norma unitaria, risulta:

$$\| (1-\lambda)x + \lambda y \| < 1 \qquad \forall \lambda \in]0,1[.$$

$$T_\varepsilon u = -A(u^* - u) - B_\varepsilon(u^* - u) + J^{-1}\left[\frac{u}{\varepsilon} - B'_\varepsilon(u^* - u)\right] \qquad \forall u \in V,$$

e consideriamo l'equazione $T_\varepsilon u = 0$, che equivale all'equazione variazionale:

(24) $\qquad u \in V: \quad \langle T_\varepsilon u, v\rangle = 0 \qquad \forall v \in V.$

Ovviamente T_ε è strettamente monotono ed emicontinuo; inoltre la condizione (1) implica che la stessa (1) è verificata se in essa si sostituisce A con T_ε. Pertanto la (24) ammette una ed una sola soluzione, diciamola w_ε, e quindi w_ε è un elemento di V tale che:

$$T_\varepsilon w_\varepsilon = 0,$$

ossia tale che:

$$-A(u^* - w_\varepsilon) - B_\varepsilon(u^* - w_\varepsilon) + J^{-1}\left[\frac{w_\varepsilon}{\varepsilon} - B'_\varepsilon(u^* - w_\varepsilon)\right] = 0.$$

Posto $u_\varepsilon = u^* - w_\varepsilon$, tale relazione si scrive:

(25) $$A u_\varepsilon + B_\varepsilon u_\varepsilon = J^{-1}\left[\frac{u^* - u_\varepsilon}{\varepsilon} - B'_\varepsilon u_\varepsilon\right],$$

cioè:

$$J\left[A u_\varepsilon + B_\varepsilon u_\varepsilon\right] = \frac{u^* - u_\varepsilon}{\varepsilon} - B'_\varepsilon u_\varepsilon,$$

ossia:

$$u_\varepsilon + \varepsilon\left[J(A + B_\varepsilon)u_\varepsilon + B'_\varepsilon u_\varepsilon\right] = u^*.$$

Pertanto u_ε soddisfa alla (6_2); d'altra parte, essendo B_ε e J^{-1} operatori a valori in W, dalla (25) si deduce che u_ε soddisfa anche alla (6_1).

* * *

Allo scopo di enunciare un teorema che assicura la validità dell'implicazione (3), introduciamo la seguente definizione.

Siano: V uno spazio di Banach riflessivo, K un sottoinsieme chiuso e convesso di V, A un operatore di V in V', monotono, emicontinuo, e soddisfacente alla (1) se K non è limitato. Sia inoltre f un elemento di V'.

Fissata una soluzione u^* della (2), diremo che la terna

$$(u^*, A, K)$$

è <u>J-compatibile</u>, quando esistono due operatori del tipo (4)

soddisfacenti alle (5) e tali che sia verificata la condizione (6). In tale ipotesi, diremo che $\{u_\varepsilon\}$ è una famiglia di elementi di K determinata da (u^*, A, K).

Orbene, detto W uno spazio di Banach riflessivo col sostegno sottospazio vettoriale di V', e con immersione continua in V', sussiste il seguente

TEOREMA. *Se la terna (u^*, A, K) è J-compatibile, si ha che:*

$$f \in W \Rightarrow Au^* \in W .$$

DIM. Nell'ipotesi che la terna (u^*, A, K) sia J-compatibile, diciamo $\{u_\varepsilon\}$ una famiglia di elementi di K da essa determinata, e supponiamo $f \in W$.

Per il teor. III esiste una sottofamiglia $\{u_{\varepsilon'}\}$ di $\{u_\varepsilon\}$ tale che:

$$\lim_{\varepsilon' \to 0} \langle Au_{\varepsilon'}, z \rangle = \langle Au^*, z \rangle \quad \forall z \in V. \tag{26}$$

D'altra parte per il teor. I la famiglia $\{Au_{\varepsilon'}\}$ è limitata in W, e quindi esiste una sottofamiglia, che denotiamo ancora con $\{Au_{\varepsilon'}\}$, debolmente convergente in W, per $\varepsilon' \to 0$, ad un elemento $\eta \in W$:

$$\lim_{\varepsilon' \to 0} \ll w', Au_{\varepsilon'} \gg = \ll w', \eta \gg \quad \forall w' \in W'. \tag{27}$$

Poiché per ogni $z \in V$ il funzionale $w \in W \subseteq V' \to \langle w, z \rangle$ è un elemento di W', dalla (27) segue:

$$\lim_{\varepsilon' \to 0} \langle Au_{\varepsilon'}, z \rangle = \langle \eta, z \rangle \quad \forall z \in V. \tag{28}$$

Da (26) e (28) si deduce:

$$\langle Au^*, z \rangle = \langle \eta, z \rangle \quad \forall z \in V,$$

ossia $Au^* = \eta$. Poiché η è un elemento di W; si conclude che $Au^* \in W$ e il teorema è dimostrato.

SOTTODIFFERENZIALE

Siano: S uno spazio vettoriale topologico localmente convesso, S^* il duale topologico di S.

Siano inoltre F un funzionale definito in S, a valori in $\overline{R}$, ed x_0 un punto di S in cui F assuma valore finito.

Se esiste un elemento u di S^* (non necessariamente unico) tale da risultare:

$$(1) \qquad F(x_0) + \langle u, x - x_0 \rangle \leq F(x) \qquad \forall x \in S,$$

si dice che u è un sottogradiente di F nel punto x_0. L'insieme dei sottogradienti di F in x_0 si chiama il sottodifferenziale di F in x_0 e si denota con $\partial F(x_0)$.

Evidentemente le proposizioni che seguono sono equivalenti:

a) u è un sottogradiente di F in x_0.

b) Il funzionale definito ponendo:

$$\Phi(x) = F(x) - \langle u, x \rangle \qquad \forall x \in S$$

è dotato di minimo, ed assume il suo minimo valore nel punto x_0.

c) Posto $k = F(x_0) - \langle u, x_0 \rangle$, la funzione affine definita ponendo:

$$\varphi(x) = \langle u, x \rangle + k \qquad \forall x \in S$$

soddisfa alle condizioni:

$$\varphi(x_0) = F(x_0), \quad \varphi(x) \leq F(x) \quad \forall x \in S \qquad (1).$$

Dall'equivalenza tra a) e b) consegue ovviamente che:

I. Condizione necessaria e sufficiente affinché il funzionale F sia dotato di minimo ed assuma il minimo valore nel punto x_0 è che si abbia:

$$0 \in \partial F(x_0),$$

cioè che 0 sia sottogradiente di F in x_0.

---

(1) L'uguaglianza si suole esprimere dicendo che la funzione affine $\varphi(x)$ è esatta in x_0. La c) si esprime allora concisamente dicendo che la funzione affine $\varphi(x)$ è minorante di $F(x)$ ed esatta in x_0.

Vogliamo ora far vedere che:

II. Se il funzionale F è convesso e differenziabile in senso debole nel punto x_0, il sottodifferenziale di F in x_0 è costituito da un unico elemento, che è il differenziale di F in x_0.

Dim. Essendo F convesso, per ogni $x \in S$ e per ogni $t \in]0,1]$ si ha:

$$F[(1-t)x_0 + tx] \leq (1-t)F(x_0) + tF(x) \quad (1)$$

ossia:

$$\frac{F[x_0 + t(x-x_0)] - F(x_0)}{t} \leq F(x) - F(x_0),$$

da cui, passando al limite per $t \to 0$, consegue [cfr. pag. 288]:

$$\langle dF(x_0), x - x_0 \rangle \leq F(x) - F(x_0).$$

Dunque l'elemento $u = dF(x_0)$ di S^* verifica la (1), ossia è un sottogradiente di F in x_0.

Per dimostrare che $dF(x_0)$ è l'unico sottogradiente di F in x_0, sia u_1 un elemento di $\partial F(x_0)$, sicché risulta:

$$F(x_0) + \langle u_1, x - x_0 \rangle \leq F(x) \quad \forall x \in S.$$

Fissato ad arbitrio $y \in S$ e posto $x = x_0 + ty$ con $t > 0$, ne segue che:

$$\langle u_1, y \rangle \leq \frac{F(x_0 + ty) - F(x_0)}{t}$$

e quindi, passando al limite per $t \to 0$:

$$\langle u_1, y \rangle \leq \langle dF(x_0), y \rangle.$$

Per l'arbitrarietà di y si ha dunque $u_1 = dF(x_0)$ e l'asserto è provato.

Notiamo che se $S = R$ ed F è una funzione reale di una variabile reale soddisfacente alle ipotesi del teor. II, la (1) con $u = dF(x_0)$ si scrive:

$$F(x_0) + F'(x_0)(x - x_0) \leq F(x) \quad \forall x \in R,$$

ed esprime che il diagramma di F è situato non al disotto della retta tangente nel punto di ascissa x_0.

---

(1) Cfr. (2) pag. 63.

Notiamo pure che se S e' uno spazio di Hilbert e $a(x,y)$ e' una forma bilineare continua in $S \times S$, per il funzionale F definito ponendo:

$$F(x) = a(x,x) \quad \forall x \in S \tag{2}$$

le ipotesi del teor. II sono verificate qualunque sia $x_0 \in S$ se risulta:

$$F(x) \geq 0 \quad \forall x \in S, \tag{3}$$

in particolare se la forma bilineare $a(x,y)$ e' coercitiva.

Infatti per la continuita' di $a(x,y)$ il funzionale definito dalla (2) e' differenziabile nel punto x_0 addirittura in senso forte [cfr. (2.3), IV pag. 291], e dalla (3) consegue facilmente che tale funzionale e' convesso.

Invero per la (3) risulta, qualunque siano $x, y \in S$:

$$a(x-y, x-y) \geq 0,$$

ossia:

$$a(x,y) + a(y,x) \leq a(x,x) + a(y,y),$$

e quindi qualunque sia $t \in [0,1]$ si ha:

$$F[(1-t)x + ty] = a\big((1-t)x+ty, (1-t)x+ty\big) = (1-t)^2 a(x,x) + t^2 a(y,y) +$$
$$+ t(1-t)\left[a(x,y) + a(y,x)\right] \leq (1-t)^2 a(x,x) + t^2 a(y,y) + t(1-t)\left[a(x,x) + a(y,y)\right] =$$
$$= (1-t)\, a(x,x) + t\, a(y,y) = (1-t) F(x) + t F(y).$$

Il teor. II in un certo senso si inverte. Precisamente si dimostra che:

III. Se il funzionale F e' convesso, e' continuo in x_0 ed ammette ivi un unico sottogradiente, allora F e' differenziabile in senso debole nel punto x_0.

UN' OSSERVAZIONE SUL TEOREMA DI VITALI-LUSIN

Il teorema di Vitali-Lusin [cfr. pag. 35] si dimostra per le funzioni definite quasi ovunque in un insieme $X \subseteq R^k$ misurabile e di misura finita. Vogliamo far vedere che la tesi sussiste anche se X ha misura infinita: precisamente, supposto che per X di misura finita sia acquisita l'equivalenza (f misurabile in X) $\Leftrightarrow$ (f quasi continua in X), vogliamo dimostrare tale equivalenza nel caso che X abbia misura infinita.

Cominciamo col dimostrare che:

$$f \text{ misurabile in } X \Rightarrow f \text{ quasi continua in } X .$$

Per ogni $n \in N$ sia C_n il cerchio aperto di R^k di raggio n. Essendo f misurabile in $X \cap C_n$, la f è ivi quasi continua: quindi fissato $\varepsilon > 0$ esiste un insieme misurabile $X_{n,\varepsilon} \subseteq X \cap C_n$, di misura minore di $\varepsilon/2^n$, tale che la restrizione di f ad $X \cap C_n - X_{n,\varepsilon}$ sia continua. Posto $X_\varepsilon = \bigcup_{n \in N} X_{n,\varepsilon}$, è continua anche la restrizione di f all'insieme $X \cap C_n - X_\varepsilon$ [perché questo è contenuto nell'insieme $X \cap C_n - X_{n,\varepsilon}$]. Siccome la misura di X_ε è minore di ε, per dimostrare l'asserto basta far vedere che la restrizione di f ad $X - X_\varepsilon$ è continua. Detto x_0 un punto di $X - X_\varepsilon$ e detto ν un intero maggiore della distanza di x_0 dall'origine, il punto x_0 appartiene ad $X \cap C_\nu - X_\varepsilon$ e perciò in tale punto è continua la restrizione di f ad $X \cap C_\nu - X_\varepsilon$: applicando la definizione di continuità, se ne deduce in modo ovvio che in x_0 è continua anche la restrizione di f ad $X - X_\varepsilon$.

Dimostriamo ora, viceversa, che se f è quasi continua in X, f è misurabile in X. Qualunque sia $a \in R$, si ha:

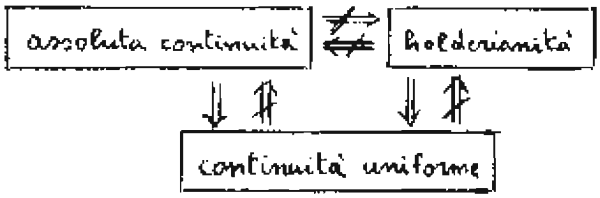
$$(*) \qquad \{x \in X : f(x) > a\} = \bigcup_{n \in N} \{x \in X \cap C_n : f(x) > a\} .$$

Essendo f quasi continua in $X \cap C_n$, l'insieme $\{x \in X \cap C_n : f(x) > a\}$ è misurabile, e perciò è misurabile anche l'insieme a primo membro della $(*)$, donde l'asserto.

CONFRONTO TRA LE NOZIONI:

HOLDERIANITA' - ASSOLUTA CONTINUITA' - CONTINUITA' UNIFORME

Schematizziamo il confronto in questione come segue.



Per dimostrare le asserzioni contenute nello schema, premettiamo alcuni esempi.

Es.1. La funzione φ definita ponendo:

$$\varphi(x)=\begin{cases}0 & \text{se } x=0\\ 1/\log x & \text{se } x\neq 0\end{cases} \qquad \forall x\in[0,\tfrac{1}{2}]$$

non è hölderiana con alcun esponente. La φ è assolutamente continua.

DIM. Ammettiamo che esistano $\lambda\in]0,1]$ e $K\in\mathbb{R}^+$ tali che $|\varphi(x')-\varphi(x'')|\le K|x'-x''|^\lambda \ \forall x\in[0,\tfrac{1}{2}]$. In particolare si avrebbe $1/|\log x|\le Kx^\lambda \ \forall x\in]0,1/2]$ e quindi $x^\lambda|\log x|\ge 1/K \ \forall x\in]0,1/2]$: disuguaglianza assurda perché il primo membro tende a 0 per $x\to 0$.
Per dimostrare che φ è assolutamente continua basta osservare che risulta:

$$\varphi(x)=\frac{1}{\log(1/2)}+\int_{1/2}^{x}\frac{-1}{t\log^2 t}\,dt \qquad \forall x\in 0,\tfrac{1}{2}]$$

e ricordare il teor. 5 di pag. 47.

Es.2. La funzione continua f definita ponendo:

$$f(x)=\begin{cases}0 & \text{se } x=0\\ x\,\mathrm{sen}\,(1/x) & \text{se } x\neq 0\end{cases} \qquad \forall x\in[0,1/\pi]$$

non è a variazione limitata, quindi non è assolutamente continua; peraltro f è assolutamente continua in ogni intervallo $[a,b]\subset]0,1/\pi]$.

La f è hölderiana con esponente $1/2$, e non è hölderiana con alcun esponente $\lambda>1/2$.

DIM. Consideriamo la partizione dell'intervallo $[0,1/\pi]$ determinata dai punti $x_i=\frac{2}{(i+1)\pi}$ con $i\in\{1,\dots,2n\}$. Per ogni $h\in\{1,\dots,n\}$ si ha $x_{2h-1}=1/h\pi$ e quindi $f(x_{2h-1})=0$; d'altra parte si ha:

$$f(x_{2h})=\frac{2}{(2h+1)\pi}>\frac{2}{(2h+2)\pi}=\frac{1}{\pi}\,\frac{1}{h+1}.$$

Risulta dunque:

$$\sum_{i=1}^{2n+1}|f(x_{i+1})-f(x_i)|\ge\sum_{h=1}^{n}|f(x_{2h})-f(x_{2h-1})|>\frac{1}{\pi}\sum_{h=1}^{n}\frac{1}{h+1},$$

sicché il primo membro è divergente per $n\to\infty$.
Pertanto f non è a variazione limitata.

Se $[a,b] \subset]0,1/\pi]$ la f è lipschiteziana in $[a,b]$, in quanto ivi dotata di derivata continua; pertanto f è assolutamente continua in $[a,b]$.

Osserviamo ora che per $x,y \in]0,1/\pi]$ con $x>y$ si ha:

$$|f(x)-f(y)| = \left|x\operatorname{sen}\tfrac{1}{x} - y\operatorname{sen}\tfrac{1}{y}\right| = \left|(x-y)\operatorname{sen}\tfrac{1}{x} + y\left(\operatorname{sen}\tfrac{1}{x}-\operatorname{sen}\tfrac{1}{y}\right)\right|$$

$$\le |x-y|^{\frac{1}{2}}\cdot|x-y|^{\frac{1}{2}} + y\left|\operatorname{sen}\tfrac{1}{x}-\operatorname{sen}\tfrac{1}{y}\right|^{\frac{1}{2}}\cdot\left|\operatorname{sen}\tfrac{1}{x}-\operatorname{sen}\tfrac{1}{y}\right|^{\frac{1}{2}} .$$

Essendo $|x-y| \le 1/\pi$ e $\left|\operatorname{sen}\tfrac{1}{x}-\operatorname{sen}\tfrac{1}{y}\right| \le 2$, dalle precedenti disuguaglianze consegue:

$$(1) \qquad |f(x)-f(y)| \le \frac{1}{\pi^{1/2}}|x-y|^{\frac{1}{2}} + 2^{1/2}\,y\left|\operatorname{sen}\tfrac{1}{x}-\operatorname{sen}\tfrac{1}{y}\right|^{\frac{1}{2}} .$$

D'altra parte, essendo la funzione seno lipschiteziana con coefficiente 1 (sicché per $t',t'' \in \mathbb{R}$ si ha $|\operatorname{sen}t'-\operatorname{sen}t''| \le |t'-t''|$), si ha:

$$y\left|\operatorname{sen}\tfrac{1}{x}-\operatorname{sen}\tfrac{1}{y}\right|^{\frac{1}{2}} \le y\left|\tfrac{1}{x}-\tfrac{1}{y}\right|^{\frac{1}{2}} = y\,\frac{|x-y|^{\frac{1}{2}}}{\sqrt{x}\sqrt{y}} = \sqrt{\frac{y}{x}}\,|x-y|^{\frac{1}{2}} < |x-y|^{\frac{1}{2}} ,$$

ed allora per la (1) risulta:

$$|f(x)-f(y)| \le \left(1/\pi^{1/2}+2^{1/2}\right)|x-y|^{\frac{1}{2}} ,$$

disuguaglianza questa che sussiste banalmente anche per $y=0$.

Pertanto la funzione f è hölderiana con esponente $1/2$.

Infine, fissato $\lambda \in]1/2,1[$, dimostriamo che f non è hölderiana con esponente λ.

Scelto un numero h soddisfacente alla relazione:

$$(2) \qquad 1 < h < \frac{\lambda}{1-\lambda} ,$$

il che è possibile in quanto $1-\lambda > 0$ e $\lambda > 1-\lambda$, poniamo:

$$x_n = \frac{1}{\pi/3+2n\pi} , \qquad y_n = t_n x_n \quad \text{con } t_n = 1-n^{-\frac{\lambda}{h(1-\lambda)}} .$$

Proveremo che la successione $\{a_n\}$ definita ponendo:

$$a_n = \frac{f(x_n)-f(y_n)}{(x_n-y_n)^\lambda} \qquad \forall n \in \mathbb{N}$$

è divergente, e da ciò seguirà ovviamente l'asserto.

Siccome risulta:

$$a_n = \frac{x_n\operatorname{sen}\frac{1}{x_n}-y_n\operatorname{sen}\frac{1}{y_n}}{(x_n-y_n)^\lambda} = \frac{(x_n-y_n)\operatorname{sen}\frac{1}{x_n}+y_n\left(\operatorname{sen}\frac{1}{x_n}-\operatorname{sen}\frac{1}{y_n}\right)}{(x_n-y_n)^\lambda} =$$

$$= \frac{\sqrt{3}}{2}(x_n-y_n)^{1-\lambda} + \frac{y_n}{(x_n-y_n)^\lambda}\left(\frac{\sqrt{3}}{2}-\operatorname{sen}\frac{1}{y_n}\right) ,$$

tenuto conto che $x_n-y_n = (1-t_n)x_n \to 0$ [$t_n \to 1$ dato che per la (2) si ha $\frac{\lambda}{h(1-\lambda)} > 1$] si ha intanto:

$$(3) \qquad \lim a_n = \lim \frac{y_n}{(x_n-y_n)^\lambda}\left(\frac{\sqrt{3}}{2}-\operatorname{sen}\frac{1}{y_n}\right) .$$

Poniamo ora:

$$\theta_n = \frac{1}{y_n}-\frac{1}{x_n} ,$$

sicché risulta:

$$(4) \qquad \frac{\sqrt{3}}{2}-\operatorname{sen}\frac{1}{y_n} = \frac{\sqrt{3}}{2}(1-\cos\theta_n)-\frac{1}{2}\operatorname{sen}\theta_n .$$

D'altra parte:

$$\theta_n = \frac{1}{t_n x_n}-\frac{1}{x_n} = \frac{1-t_n}{t_n x_n} = \frac{1}{t_n}\cdot\frac{1}{x_n}(1-t_n) = \frac{1}{t_n}\left(\frac{\pi}{3}+2n\pi\right)(1-t_n) =$$

$$= \frac{1}{t_n}\,\frac{\pi/3+2n\pi}{n}\cdot n^{1-\frac{\lambda}{h(1-\lambda)}}$$

e quindi $\theta_n \to 0$. Tenuto conto che

$$\operatorname{sen}\frac{1}{y_n} = \operatorname{sen}\left(\theta_n + \pi/3 + 2n\pi\right) = \operatorname{sen}\left(\theta_n + \frac{\pi}{3}\right)$$

da cui:

$$\frac{\sqrt{3}}{2} - \operatorname{sen}\frac{1}{y_n} = \frac{\sqrt{3}}{2}(1-\cos\theta_n) - \frac{1}{2}\operatorname{sen}\theta_n ,$$

dalla (4) segue:

$$\lim \frac{1}{\theta_n}\left(\frac{\sqrt{3}}{2} - \operatorname{sen}\frac{1}{y_n}\right) = -\frac{1}{2} .$$

Per la (3) si ha quindi:

(5) $$\lim a_n = -\frac{1}{2}\lim \frac{y_n\theta_n}{(x_n-y_n)^\lambda} .$$

E' facile ora calcolare il limite della successione (a_n), tenendo presente che per la definizione di θ_n si ha $\theta_n y_n = 1 - y_n/x_n = 1-t_n$, e che $x_n - y_n = x_n - x_n t_n = x_n(1-t_n)$. Si ha dunque:

$$\frac{y_n\theta_n}{(x_n-y_n)^\lambda} = \frac{1}{x_n^\lambda}(1-t_n)^{1-\lambda} = \frac{1}{x_n^\lambda} n^{-\frac{\lambda}{h}} = \left(\frac{\pi/3+2n\pi}{n}\right)^\lambda n^\lambda \cdot n^{-\frac{\lambda}{h}} = \left(\frac{\pi/3+2n\pi}{n}\right)^\lambda n^{\lambda\left(1-\frac{1}{h}\right)}$$

ed allora dalla (5), dato che $\frac{1}{h} < 1$ per la (2), si deduce che $\lim a_n = -\infty$.

Es. 3. A complemento dell'es. 2 osserviamo che anche la funzione continua g definita ponendo:

$$g(x) = \begin{cases} 0 & \text{se } x = 0 \\ x\operatorname{sen}^2(\pi/\sqrt{x}) & \text{se } x \neq 0 \end{cases} \qquad \forall x \in [0,1]$$

è hölderiana con esponente 1/2.

Di conseguenza la funzione $\sqrt{g}$, non assolutamente continua, è hölderiana con esponente 1/4.

Dim. Siano x, y punti di $]0,1]$, con $x > y$. Mostriamo preliminarmente che la funzione $\sqrt{t}$ è hölderiana con esponente 1/2, risultando:

(6) $$\sqrt{x} - \sqrt{y} \leq (x-y)^{1/2} .$$

Alla (6) si perviene scrivendo il primo membro nella forma $\int_y^x (1/2\sqrt{t})\,dt$, e maggiorando la funzione sotto il segno di integrale con $1/2\sqrt{t-y}$ [dato che $t > t-y$]. Si ottiene così:

$$\sqrt{x} - \sqrt{y} < \int_y^x \frac{1}{2\sqrt{t-y}}\,dt = (x-y)^{1/2} .$$

Ciò premesso, tenuto conto che la funzione $\operatorname{sen}^2 t$ è lipschitziana con coefficiente 2, si ha:

$$|g(x) - g(y)| = \left|x\operatorname{sen}^2\frac{\pi}{\sqrt{x}} - y\operatorname{sen}^2\frac{\pi}{\sqrt{y}}\right| = \left|(x-y)\operatorname{sen}^2\frac{\pi}{\sqrt{x}} + y\left(\operatorname{sen}^2\frac{\pi}{\sqrt{x}} - \operatorname{sen}^2\frac{\pi}{\sqrt{y}}\right)\right| \leq$$

$$\leq (x-y) + 2\pi y\left|\frac{1}{\sqrt{x}} - \frac{1}{\sqrt{y}}\right| \leq (x-y)^{1/2} + 2\pi y\frac{(\sqrt{x}-\sqrt{y})}{\sqrt{xy}} \leq (x-y)^{1/2} + 2\pi\sqrt{\frac{y}{x}}(\sqrt{x}-\sqrt{y})$$

$$\leq (x-y)^{1/2} + 2\pi(x-y)^{1/2} = (2\pi+1)|x-y|^{1/2}$$

e pertanto g è hölderiana con esponente 1/2.

Il fatto che $\sqrt{g}$ non è assolutamente continua si prova con procedimento analogo a quello seguito per la funzione f dell'esempio 2.

Dimostriamo ora le relazioni schematizzate all'inizio.

a) Le nozioni di assoluta continuità e di hölderianità sono indipendenti: la prima non implica la seconda per l'es. 1, la seconda non implica la prima per l'es. 2.

b) L'assoluta continuità implica la continuità uniforme, non viceversa.

Se f è una funzione assolutamente continua in $[a,b]$, dalla definizione di assoluta continuità [pag. 47] si deduce in particolare che: per ogni $\varepsilon>0$ esiste un $\delta_\varepsilon>0$ tale che, qualunque sia $[x',x'']\subseteq[a,b]$, si abbia $x''-x'<\delta_\varepsilon \Rightarrow |f(x'')-f(x')|<\varepsilon$; in altri termini, si deduce che f è uniformemente continua. Perciò l'implicazione da dimostrare è banale.

L'implicazione opposta non sussiste: basta considerare la funzione dell'es. 2, che è uniformemente continua in quanto è continua in un intervallo compatto.

c) L'hölderianità implica la continuità uniforme, non viceversa.

Per la stessa definizione di continuità uniforme, le funzioni hölderiane sono uniformemente continue. Per l'es. 1, esistono funzioni uniformemente continue non hölderiane.

Osservazioni.

1. Ogni funzione f_0 uniformemente continua in un aperto Ω è prolungabile per continuità su $\partial\Omega$, e il prolungamento f è una funzione uniformemente continua in $\overline{\Omega}$.

DIM. Se $\bar{x}$ è un punto di $\partial\Omega$, $\bar{x}$ è di accumulazione per Ω e, in virtù della continuità uniforme, f_0 soddisfa alla condizione del criterio di convergenza di Cauchy: pertanto f_0 è prolungabile per continuità in $\bar{x}$.

Sia f il prolungamento continuo di f_0 su $\partial\Omega$. Se Ω è limitato, la continuità uniforme di f in $\overline{\Omega}$ scaturisce dal teorema di Heine-Cantor; il ragionamento che segue si riferisce perciò al caso Ω non limitato.

Fissato $\varepsilon>0$, per l'ipotesi esiste $\delta_\varepsilon>0$ tale che:

(1) $\qquad x',x''\in\Omega \ \text{e}\ |x'-x''|<\delta_\varepsilon \Rightarrow |f_0(x')-f_0(x'')|<\varepsilon/4$.

Vogliamo mostrare che, fissati due punti qualunque x',x'' di $\overline{\Omega}$ soddisfacenti alla relazione

(2) $\qquad |x'-x''|<\delta_\varepsilon/2$,

risulta:

(3) $\qquad |f(x')-f(x'')|<\varepsilon$.

Se x' ed x'' appartengono entrambi ad Ω, dalla (2) consegue la (3) in virtù di (1).

Supponiamo che sia $x'\in\partial\Omega$, $x''\in\Omega$. Siccome $f(x')=\lim_{x\to x'} f_0(x)$, esiste $\sigma_\varepsilon(x')<\delta_\varepsilon/2$

tale che:

$$x \in \Omega \ , \ |x-x'| < \sigma_\varepsilon(x') \Rightarrow |f(x')-f_0(x)| < \varepsilon/4 \ .$$

Nel cerchio aperto di centro x' e raggio $\sigma_\varepsilon(x')$ scegliamo allora un punto $x \in \Omega$, sicché risulta $|f(x')-f(x)| < \varepsilon/4$; d'altra parte ricordando la (2) si ha $|x-x''| \le |x-x'|+|x'-x''| < \sigma_\varepsilon(x') + \delta_\varepsilon/2 < \delta_\varepsilon/2 + \delta_\varepsilon/2 = \delta_\varepsilon$, e quindi per la (1) risulta $|f(x)-f(x'')| < \varepsilon/4$. Ne consegue che $|f(x')-f(x'')| \le |f(x')-f(x)| + |f(x)-f(x'')| < \varepsilon/2$, e quindi la (3) è vera anche nel caso in esame. Notiamo che il ragionamento testè concluso prova anche l'implicazione:

(4) $\qquad x \in \Omega \ , \ y \in \partial\Omega \ e \ |x-y| < \delta_\varepsilon/2 \Rightarrow |f(x)-f(y)| < \varepsilon/2 \ .$

Esaminiamo infine il caso $x' \in \partial\Omega$, $x'' \in \partial\Omega$.

Detto C' il cerchio aperto di centro x' e raggio $\delta_\varepsilon/2$, consideriamo un cerchio $C'' \subset C'$ di centro x'', il cui raggio è certamente minore di $\delta_\varepsilon/2$. Scelto in C'' un punto $x \in \Omega$, risultando $|x-x''| < \delta_\varepsilon/2$ e $|x-x'| < \delta_\varepsilon/2$ si ha per la (4) che $|f(x)-f(x'')| < \varepsilon/2$ e $|f(x)-f(x')| < \varepsilon/2$: ne segue $|f(x')-f(x'')| \le |f(x')-f(x)| + |f(x)-f(x'')| < \varepsilon$ e quindi ancora la (3).

Da quanto stabilito si deduce che f è uniformemente continua, ossia la tesi.

— — — — — — —

2. Ogni funzione f_0 hölderiana con esponente λ in un aperto Ω è prolungabile per continuità su $\partial\Omega$. Il prolungamento f è una funzione hölderiana in $\overline{\Omega}$ con lo stesso esponente λ e con lo stesso coefficiente.

— — — — — —

DIM. La prima asserzione consegue dalla 1, ricordando la proposizione c).

Per dimostrare la seconda asserzione, consideriamo la disuguaglianza:

(5) $\qquad |f_0(x')-f_0(x'')| \le [f_0]_\lambda |x'-x''|^\lambda$

valida qualunque siano $x', x'' \in \Omega$. Fissati $x, y \in \overline{\Omega}$, la disuguaglianza

(6) $\qquad |f(x)-f(y)| \le [f_0]_\lambda |x-y|^\lambda$

è banale per la (5) se x ed y appartengono entrambi ad Ω. Se uno solo dei due punti appartiene ad Ω, ad esempio x, dalla (5) con $x' = x$ passando al limite per $x'' \to y$ si deduce la (6). Se x ed y appartengono entrambi a $\partial\Omega$, dalla (5) passando al limite per $x' \to x$ si deduce che la (5) sussiste con $x' = x$; passando poi al limite per $x'' \to y$ si ottiene la (6).

— — — — — —

In virtù della 2, si può dire che a meno di restrizioni e prolungamenti gli spazi $C^{0,\lambda}(\Omega)$ e $C^{0,\lambda}(\overline{\Omega})$ coincidono.

3. Una funzione continua in $[c,d]$ e assolutamente continua in ogni intervallo compatto contenuto in $]c,d[$, non necessariamente è assolutamente continua in $[c,d]$.

Basta tener presente l'esempio 2 di pag. 330.

Nelle pagine successive alla presente riprenderemo gli operatori integrali introdotti a pag. 169, al fine di riportare alcuni risultati relativi alle equazioni integrali di Fredholm.

Per motivi che qui è superfluo spiegare, non abbiamo potuto proseguire la numerazione delle pagine, che saranno invece numerate da 1 a 16 (1).

L'argomento in questione sarà preceduto da una "Introduzione", anche questa con numerazione autonoma delle pagine (da 1 a 16).

Successivamente, ancora con numerazione autonoma delle pagine, ci occuperemo di una generalizzazione del concetto di integrale superficiale, che conduce ad una formulazione molto generale del teorema della divergenza e del teorema di Stokes. Questo argomento è stato tratto dal testo "Analisi Matematica due" di Fusco-Marcellini-Sbordone.

Infine daremo un esempio di funzione continua in un intervallo, con derivata nulla quasi ovunque, che non è costante.

---

(1) A pag. 9 verrà citato un teorema che riportiamo qui.

Teorema di Beppo Levi. Sia $(f_n)_{n\in\mathbb{N}}$ una successione q.o. crescente di funzioni non negative e sommabili in un aperto Ω di $\mathbb{R}^m$.

Se la successione degli integrali delle f_n è limitata, si ha:

$$\lim_n \int_\Omega f_n(t)\,dt = \int_\Omega (\lim f_n)(t)\,dt .$$

Nella predetta pag. 9 il teorema si applica alla successione definita ponendo, per ogni $n\in\mathbb{N}$ e per s quasi ovunque in Ω:

$$f_n(s) = \sum_{i=1}^{n} |\lambda_i|^2 |x_i(s)|^2 .$$

La successione degli integrali è limitata perché maggiorata da $\|K\|^2$, essendo K il nucleo dell'operatore integrale in questione.

Introduzione

alle equazioni integrali di Fredholm.

Sia $M=\{a_{ij}\}$ una matrice ad elementi complessi, con infinite righe ed infinite colonne. Poniamo:

$$\|M\| = \lim_{n} \left(\sum_{i,j}^{1\dots n} |a_{ij}|^2 \right)^{\frac{1}{2}}$$

supponendo che il limite, ovviamente esistente, sia finito e non nullo.

Occorrerà considerare lo spazio normato l_2 costituito dalle successioni $\sigma=(x_n)_{n\in N}$ di numeri complessi che siano di quadrato sommabile, cioè tali che la serie $\sum_{n\in N} |x_n|^2$ sia convergente; la norma di σ è definita ponendo:

$$\|\sigma\|_{l_2} = \left(\sum_{n\in N} |x_n|^2 \right)^{\frac{1}{2}}.$$

Per ogni siffatta successione, qualunque sia $i\in N$ la serie:

$$\sum_{j=1}^{\infty} a_{ij} x_j$$

è assolutamente convergente, quindi convergente. Infatti per la somma parziale n^{ma} della serie dei moduli si ha:

$$\sum_{j=1}^{n} |a_{ij} x_j| \le \left(\sum_{j=1}^{n} |a_{ij}|^2 \right)^{1/2} \left(\sum_{j=1}^{n} |x_j|^2 \right)^{1/2} \le \left(\sum_{i,j}^{1\dots n} |a_{ij}|^2 \right)^{1/2} \|\sigma\| \le \|M\| \cdot \|\sigma\| .$$

Pertanto, assegnata la successione $(y_n)_{n\in N} \in l_2$, ha senso considerare il sistema nell'incognita $(x_n)_{n\in N} \in l_2$:

$$(1) \qquad \begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n + \dots = y_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n + \dots = y_2 \\ \dots\dots\dots\dots\dots\dots\dots \\ a_{n1}x_1 + a_{n2}x_2 + \dots + a_{nn}x_n + \dots = y_n \\ \dots\dots\dots\dots\dots\dots\dots \end{cases}$$

che può chiamarsi, impropriamente ma efficacemente, un sistema lineare di infinite equazioni in infinite incognite.

Ci proponiamo di mostrare che il problema espresso dalle (1), ambientato nello spazio l_2, è equivalente ad un'equazione ambientata in un opportuno sottospazio S di $L^2(\Omega)$,

con

$$\Omega = \bigcup_{i \in N}]i, i+1[.$$

Le equazioni integrali di Fredholm, che sono ambientate nello spazio $L^2(\Omega)$ con Ω aperto di R^k, possono considerarsi una generalizzazione di quella che equivale al sistema (1).

Il sottospazio S che prenderemo in considerazione e' il codominio dell'operatore lineare

$$B : \ell_2 \longrightarrow L^2(\Omega)$$

definito come segue. Per ogni successione

$$\sigma = (x_n)_{n \in N} \in \ell_2 ,$$

denotiamo con x la funzione costante in ciascun intervallo $]i, i+1[$, con $i \in N$, definita ponendo:

$$x(s) = x_i \qquad \forall s \in]i, i+1[. \tag{2}$$

Tale funzione x rappresenta un elemento di $L^2(\Omega)$. Infatti per ogni $t \in [1, +\infty[$ si ha, qualunque sia $n > t$:

$$\int_1^t |x(s)|^2 ds \le \int_1^{n+1} |x(s)|^2 ds = \sum_{j=1}^{n} \int_j^{j+1} |x(s)|^2 ds = \sum_{j=1}^{n} |x_j|^2 \le \sum_{j=1}^{\infty} |x_j|^2$$

donde la sommabilita' di $|x|^2$ in $]1, +\infty[$, quindi in Ω.

Orbene, $B\sigma$ e' l'elemento di $L^2(\Omega)$ che ha come rappresentante la predetta funzione x. Essendo l'operatore B ovviamente lineare, il suo codominio S e' un sottospazio di $L^2(\Omega)$ (1).

Posto:

$$\Omega_{ij} =]i, i+1[\times]j, j+1[\qquad \forall i, j \in N ,$$

consideriamo ora la funzione K, costante in ciascun intervallo Ω_{ij}, definita in $\Omega \times \Omega$ ponendo:

$$K(s,t) = a_{ij} \qquad \forall (s,t) \in \Omega_{ij} \tag{3}$$

Introduciamo quindi l'operatore lineare:

$$J : L^2(\Omega) \longrightarrow L^2(\Omega)$$

che ad ogni $x \in L^2(\Omega)$ associa la funzione definita in Ω ponendo:

$$Jx(s) = \int_\Omega K(s,t) x(t) dt \qquad \forall s \in \Omega . \tag{4}$$

---

(1) Osserviamo che il prodotto scalare di due elementi di ℓ_2, diciamoli $(\alpha_n)_{n \in N}$ e $(\beta_n)_{n \in N}$, coincide con il prodotto scalare delle loro immagini mediante B, diciamole α e β. Si ha infatti:

$$(\alpha, \beta)_{L^2(\Omega)} = \int_\Omega \alpha(s) \overline{\beta}(s) ds = \sum_{j=1}^{\infty} \alpha_j \overline{\beta}_j = \left((\alpha_n)_{n \in N}, (\beta_n)_{n \in N} \right)_{\ell_2} .$$

Di conseguenza si ha:

$$\| B\sigma \|_{L^2(\Omega)} = \| \sigma \|_{\ell_2} \qquad \forall \sigma \in \ell_2 ,$$

sicche' l'operatore B e' un isomorfismo isometrico [pag. 139] di ℓ_2 sul sottospazio S codominio di B.

Occorre dimostrare che è lecito considerare l'operatore J : cioè in primo luogo che per ogni $s \in \Omega$ ha significato l'integrale a secondo membro della (4), e poi che la funzione Jx definita mediante la (4) rappresenta un elemento di $L^2(\Omega)$.

— — — — — — — — — —

DIM. Fissato $s \in \Omega$, supponiamo che sia $s \in]i, i+1[$. Per provare che la funzione:

$$t \in \Omega \to K(s,t)\,x(t)$$

è sommabile in Ω, essendo $x \in L^2(\Omega)$ basta far vedere che la funzione $t \in \Omega \to K(s,t)$ rappresenta un elemento di $L^2(\Omega)$, e ciò consegue dalla relazione, valida per ogni $n \in N$:

$$\int_1^{n+1} |K(s,t)|^2 dt = \sum_{j=1}^{n} \int_j^{j+1} |K(s,t)|^2 dt = \sum_{j=1}^{n} |a_{ij}|^2 dt \le \|M\|^2 .$$

Al fine di dimostrare che la funzione Jx definita dalla (4) è di quadrato sommabile in Ω, osserviamo intanto che per la disuguaglianza di Schwarz si ha:

$$|Jx(s)|^2 = \left| \int_\Omega K(s,t)x(t)\,dt \right|^2 \le \left(\int_\Omega |K(s,t)| \cdot |x(t)|\,dt \right)^2 \le \int_\Omega |K(s,t)|^2 dt \cdot \int_\Omega |x(t)|^2 dt \quad \forall s \in \Omega .$$

Tale relazione, posto:

$$\varphi(s) = \int_\Omega |K(s,t)|^2 dt \qquad \forall s \in \Omega ,$$

diviene:

$$(5) \qquad |Jx(s)|^2 \le \varphi(s) \cdot \|x\|^2 \quad \forall s \in \Omega .$$

Orbene, proviamo che la funzione φ è sommabile in Ω : dopo ciò, dalla (5) seguirà l'asserto. Tenuto conto che la serie $\sum_{j \in N} |a_{ij}|^2$ è convergente, in quanto le somme parziali sono maggiorate da $\|M\|^2$, per ogni $s \in]i, i+1[$ si ha:

$$\varphi(s) = \int_\Omega |K(s,t)|^2 dt = \lim_n \int_1^{n+1} |K(s,t)|^2 dt = \lim_n \sum_{j=1}^{n} \int_j^{j+1} |K(s,t)|^2 dt = \lim_n \sum_{j=1}^{n} |a_{ij}|^2 = \sum_{j=1}^{\infty} |a_{ij}|^2$$

e quindi per ogni $n \in N$ risulta:

$$\int_1^{n+1} \varphi(s)\,ds = \sum_{i=1}^{n} \int_i^{i+1} \varphi(s)\,ds = \sum_{i=1}^{n} \int_i^{i+1} \sum_{j=1}^{\infty} |a_{ij}|^2 ds = \sum_{i=1}^{n} \sum_{j=1}^{\infty} |a_{ij}|^2 \le \|M\|^2$$

donde la sommabilità di φ in $]1, +\infty[$, quindi in Ω.

— — — — — — — — — —

E' importante per i nostri scopi stabilire che risulta $J(S) \subseteq S$, cioè che sussiste l'implicazione:

$$(6) \qquad x \in S \Rightarrow Jx \in S .$$

— — — — — — — — — —

DIM. Se x appartiene al sottospazio S codominio dell'operatore B, x è l'immagine mediante B di una successione $(x_n)_{n \in N} \in l_2$, cioè per ciascun intervallo $]j, j+1[$ risulta:

$$x(t) = x_j \qquad \forall t \in]j, j+1[.$$

Si ha allora:

$$\int_1^{n+1} K(s,t)x(t)\,dt = \sum_{j=1}^{n} \int_j^{j+1} K(s,t)x(t)\,dt = \sum_{j=1}^{n} x_j \int_j^{j+1} K(s,t)\,dt$$

e quindi per $s \in]i, i+1[$ risulta, ricordando la (3):

$$\int_1^{n+1} K(s,t)x(t)\,dt = \sum_{j=1}^{n} a_{ij} x_j .$$

Detto y_i il limite del secondo membro per $n \to \infty$, ne segue:

$$(7) \qquad Jx(s) = y_i \qquad \forall s \in]i, i+1[$$

sicché la funzione Jx definita dalla (4) è costante in ciascun intervallo $]i, i+1[$. Si ha poi:

$$\sum_{i=1}^{n} |y_i|^2 = \sum_{i=1}^{n} \left| \int_i^{i+1} y_i \, ds \right|^2 = \sum_{i=1}^{n} \left| \int_i^{i+1} Jx(s) \, ds \right|^2$$

$$\leq \sum_{i=1}^{n} \left(\int_i^{i+1} 1 \cdot |Jx(s)| \, ds \right)^2 \leq \sum_{i=1}^{n} \int_i^{i+1} 1^2 ds \cdot \int_i^{i+1} |Jx(s)|^2 ds = \int_1^{n+1} |Jx(s)|^2 ds \leq \|Jx\|^2 .$$

Dunque la successione $t = (y_n)_{n \in \mathbb{N}}$ è un elemento di l_2. Siccome per la (7) $Bt = Jx$, si conclude che Jx appartiene al codominio di B, ossia al sottospazio S.

— — — — — — — — — —

In definitiva, consideriamo l'equazione ambientata nel sottospazio S:

$$(8) \qquad Jx = y ,$$

cioè il problema di ricercare, assegnato $y \in S$, un elemento x di S soddisfacente alla (8).

Sussiste il seguente

TEOREMA. Siano $(x_n)_{n \in \mathbb{N}}$, $(y_n)_{n \in \mathbb{N}}$ elementi di l_2, ed x, y le rispettive immagini mediante l'operatore B, cioè le funzioni costanti a tratti definite ponendo, per $i, j \in \mathbb{N}$:

$$(9) \qquad x(t) = x_j \quad \forall t \in]j, j+1[\quad , \quad y(s) = y_i \quad \forall s \in]i, i+1[.$$

Le proposizioni che seguono sono equivalenti:

a) La successione $(x_n)_{n \in \mathbb{N}}$ è soluzione del sistema (1).

b) L'elemento x di S è soluzione dell'equazione (8).

— — — — — — — — — —

DIM. Supposto che sia vera la a), in ciascun intervallo $]i, i+1[$ risulta, per ogni $s \in]i, i+1[$ [cfr. (4)]:

$$Jx(s) = \int_\Omega K(s,t) x(t) \, dt = \lim_n \int_1^{n+1} K(s,t) x(t) \, dt = \lim_n \sum_{j=1}^{n} \int_j^{j+1} K(s,t) x(t) \, dt$$

cioè, per le (3), (9):

$$Jx(s) = \lim_n \sum_{j=1}^{n} a_{ij} x_j = \sum_{j=1}^{\infty} a_{ij} x_j = y_i = y(s) \qquad \forall s \in]i, i+1[.$$

Dunque risulta $Jx = y$, e quindi è vera la b).

Viceversa, supponiamo x soluzione della (8). Fissati $i \in \mathbb{N}$ ed $s \in]i, i+1[$, essendo per la (3):

$$\int_j^{j+1} K(s,t) = a_{ij} ,$$

tenuto conto delle (9) si ha:

$$\sum_{j=1}^{n} a_{ij} x_j = \sum_{j=1}^{n} \int_j^{j+1} K(s,t) x_j \, dt = \sum_{j=1}^{n} \int_j^{j+1} K(s,t) x(t) \, dt = \int_1^{n+1} K(s,t) x(t) \, dt$$

e quindi:

$$\sum_{j=1}^{\infty} a_{ij} x_j = \int_\Omega K(s,t) x(t) \, dt = Jx(s) = y(s) = y_i .$$

Pertanto la successione $(x_n)_{n \in \mathbb{N}}$ è soluzione del sistema (1), ossia è vera la a).

— — — — — — — — — —

Il problema espresso dal sistema (1), formulato mediante l'equazione (8), è suscettibile di un'ampia generalizzazione.

Intanto si può porre il problema più generale di ricercare, assegnato $y \in L^2(\Omega)$, un elemento x di $L^2(\Omega)$ soddisfacente alla (8), non necessariamente dotato di un rappresentante costante a tratti. La questione si generalizza ulteriormente introducendo un parametro complesso λ e considerando l'equazione nell'incognita $x \in L^2(\Omega)$:

$$Jx - \lambda x = y \,. \tag{10}$$

Orbene, l'operatore J è un particolare operatore integrale, e la (10) è una particolare equazione integrale di Fredholm.

In generale, all'insieme $\Omega \subseteq R$ da noi considerato subentra un qualunque aperto $\Omega \subseteq R^k$, e alla funzione $K(s,t)$ utilizzata in questa "Introduzione" subentra un qualunque elemento dello spazio $L^2(\Omega \times \Omega)$, chiamato il nucleo dell'operatore integrale J (da non confondere con $\ker J$, che è il sottospazio di $L^2(\Omega)$ costituito dagli x tali che $Jx = 0$)

Naturalmente, per poter considerare l'operatore J (che si dice generato dal nucleo K) occorre dimostrare, come abbiamo fatto nel caso particolare da noi esaminato, che per ogni $x \in L^2(\Omega)$ ha senso l'integrale a secondo membro della (4), e poi che la funzione definita dalla (4) rappresenta un elemento dello spazio $L^2(\Omega)$.

* * *

Prima di affrontare lo studio delle equazioni integrali, ci sembra utile esporre qualche esempio semplice in cui l'equazione si può studiare prescindendo dalla teoria generale.

Premettiamo la seguente

OSSERVAZIONE. Sia $\Omega =]a,b[$ un intervallo limitato di R. Comunque si assegnino $h_1, h_2 \in R$, esistono infiniti elementi x dello spazio $L^2(\Omega)$ tali che:

$$\int_a^b x(t)\,dt = h_0 \quad , \quad \int_a^b t\,x(t)\,dt = h_1 \,. \tag{11}$$

DIM. Posto $x(t) = a_0 + a_1 t + a_2 t^2 \;\; \forall t \in [a,b]$, ricerchiamo i numeri a_0, a_1, a_2 in modo da verificare le (11). Si ottiene un sistema lineare di due equazioni nelle incognite a_0, a_1, a_2, e accade che il rango della matrice dei coefficienti è 2: perciò il sistema ammette infinite soluzioni.

Stante l'oggetto dei nostri discorsi, è utile evidenziare che le (11) costituiscono un sistema di due equazioni integrali: nella prima l'operatore integrale è generato dal nucleo $K(s,t) = 1$, nella seconda dal nucleo $K(s,t) = t$.

Negli esempi 1 e 2 supporremo che $\Omega =]a,b[$ sia un intervallo limitato di R.
Fissati $u, v, y \in L^2(\Omega)$ e $\lambda \in \mathbb{C}$, prenderemo in considerazione l'equazione integrale:

$$\int_a^b [u(s)+tv(s)]\,x(t)\,dt - \lambda x(s) = y(s),$$

che nell'esempio 1 sarà studiata nel caso particolare

$$u = \alpha w \quad , \quad v = \beta w$$

con $\alpha, \beta \in \mathbb{C}$ e $w \in L^2(\Omega)$.

Un altro caso particolare [$u(s)=s$, $v(s)=1$] è riportato a pag. 237 e ripreso successivamente a pag. 251.

Nell'esempio 3 studieremo un sistema lineare di infinite equazioni in infinite incognite.

ESEMPIO 1.

Fissati i numeri complessi α, β non entrambi nulli e un elemento $w \neq 0$ dello spazio $L^2(\Omega)$, consideriamo per ogni $\lambda \in \mathbb{C}$ l'equazione integrale nell'incognita $x \in L^2(\Omega)$:

$$\int_a^b w(s)(\alpha+\beta t)\,x(t)\,dt - \lambda x(s) = y(s) \qquad \text{q.o. in } \Omega \tag{12}$$

dove y è un assegnato elemento di $L^2(\Omega)$.

L'operatore integrale J è generato dal nucleo

$$K(s,t) = w(s)(\alpha+\beta t)$$

ed è definito ponendo, per ogni $x \in L^2(\Omega)$:

$$Jx(s) = \int_a^b w(s)(\alpha+\beta t)\,x(t)\,dt \qquad \text{q.o. in } \Omega. \tag{13}$$

Introduciamo le seguenti notazioni:

$$\begin{cases} w_0 = \int_a^b w(t)\,dt \\ w_1 = \int_a^b t\,w(t)\,dt \end{cases} \qquad \begin{cases} y_0 = \int_a^b y(t)\,dt \\ y_1 = \int_a^b t\,y(t)\,dt \end{cases} \tag{14}$$

che rappresentano costanti assegnate, in quanto determinate dai dati del problema.

Per lo studio della (12) esaminiamo dapprima il caso $\lambda \neq 0$.

Se $x \in L^2(\Omega)$ è un'eventuale soluzione della (12), posto:

$$c = \int_a^b (\alpha+\beta t)\,x(t)\,dt \tag{15}$$

si ha:

$$c\,w(s) - \lambda x(s) = y(s) \qquad \text{q.o. in } \Omega$$

e quindi:

$$x = \frac{1}{\lambda}(cw - y). \tag{16}$$

Siccome x deve verificare la (12), ne segue:

$$\frac{w(s)}{\lambda}\int_a^b(\alpha+\beta t)\,[c\,w(t)-y(t)]\,dt-[c\,w(s)-y(s)]=y(s)\quad \text{q.o. in } \Omega$$

ossia:

$$w(s)\int_a^b(\alpha+\beta t)\,[c\,w(t)-y(t)]\,dt-\lambda c\,w(s)=0\quad \text{q.o. in } \Omega.$$

Utilizzando le notazioni (14), tale uguaglianza diviene:

$$w(s)\,(\alpha c w_0+\beta c w_1-\alpha y_0-\beta y_1-\lambda c)=0\quad \text{q.o. in } \Omega$$

e da questa si deduce, essendo per ipotesi $w\neq 0$:

$$c\,(\alpha w_0+\beta w_1-\lambda)=\alpha y_0+\beta y_1 \qquad (17)$$

Dunque se x è un'eventuale soluzione della (12), necessariamente x ha l'espressione (16) con c soddisfacente alla (17).

Viceversa con facili calcoli si prova che, se x è l'elemento di $L^2(\Omega)$ espresso dalla (16) con la costante c soddisfacente alla (17), l'integrale a secondo membro della (15) è uguale alla costante c, e x è soluzione della (12). Da ciò si traggono, per l'equazione (12) con $\lambda\neq 0$, le seguenti conclusioni.

Se $\lambda\neq\alpha w_0+\beta w_1$ esiste un un'unica costante c soddisfacente alla (17), e quindi l'equazione integrale (12) ammette una e una sola soluzione, data dalla (16) con

$$c=\frac{\alpha y_0+\beta y_1}{\alpha w_0+\beta w_1-\lambda} \qquad (18)$$

Se invece $\lambda=\alpha w_0+\beta w_1$, e se $\alpha y_0+\beta y_1=0$, l'equazione integrale (12) ammette infinite soluzioni, date dalla (16) con c costante arbitraria: infatti la (17) diviene $c\cdot 0=0$ e quindi è soddisfatta da qualunque costante c.

Ancora per $\lambda=\alpha w_0+\beta w_1$, se risulta $\alpha y_0+\beta y_1\neq 0$ la (17) non è soddisfatta da alcun valore di c, e quindi l'equazione integrale (12) non ammette soluzioni.

Passiamo ora allo studio della (12) con $\lambda=0$, cioè dell'equazione integrale:

$$\int_a^b w(s)(\alpha+\beta t)x(t)\,dt=y(s)\quad \text{q.o. in } \Omega$$

che si può scrivere:

$$w(s)\left[\alpha\int_a^b x(t)\,dt+\beta\int_a^b t\,x(t)\,dt\right]=y(s)\quad \text{q.o. in } \Omega. \qquad (19)$$

Siccome il primo membro è proporzionale a w, è chiaro che se il secondo membro y non è proporzionale a w l'equazione (19) non ammette soluzioni

Supponiamo allora che sia $y=kw$ con $k\in\mathbb{C}$. Tenuto conto dell'ipotesi $w\neq 0$ la (19) diviene:

$$\alpha\int_a^b x(t)\,dt+\beta\int_a^b t\,x(t)\,dt=k. \qquad (20)$$

E' facile constatare che la (20) ammette infinite soluzioni. Infatti esistono

infinite coppie (h_0, h_1) di numeri complessi tali che:

$$\alpha h_0 + \beta h_1 = k \ ,$$

e per ciascuna di tali coppie esistono infiniti elementi x di $L^2(\Omega)$ tali da risultare [Oss. pag.5]:

$$\int_a^b x(t)\,dt = h_0 \ , \quad \int_a^b t\,x(t)\,dt = h_1 \ .$$

Ovviamente i predetti elementi di $L^2(\Omega)$ sono soluzioni della (20).

Teoria spettrale per l'operatore J [cfr. pag.239].

Detto I l'operatore identico di $L^2(\Omega)$ in $L^2(\Omega)$ e tenuto conto della (13), l'equazione integrale (12) si può scrivere:

$$(J - \lambda I)\,x = y \ .$$

Mostriamo che ciascun numero complesso λ, non appartenente all'insieme $\{0, \alpha w_0 + \beta w_1\}$, appartiene all'insieme risolvente dell'operatore J.

Come abbiamo visto, se $\lambda \neq 0$ e $\lambda \neq \alpha w_0 + \beta w_1$ l'equazione ammette una e una soluzione qualunque sia il termine noto $y \in L^2(\Omega)$: ciò vuol dire che l'operatore $J - \lambda I$ è invertibile ed ha per codominio tutto lo spazio $L^2(\Omega)$. Per provare che λ appartiene all'insieme risolvente di J basta allora far vedere [n.2 pag.240] che J è continuo, ossia [(2.1) pag.138] che esiste una costante k tale da risultare:

$$\|Jx\| \leq k\|x\| \quad \forall x \in L^2(\Omega).$$

Tenendo presente la disuguaglianza di Schwarz, dalla (13) si deduce:

$$|Jx(s)| \leq |w(s)| \int_a^b |\alpha + \beta t| \cdot |x(t)|\,dt \leq |w(s)| \left(\int_a^b |\alpha + \beta t|^2 dt \right)^{1/2} \|x\|$$

e quindi:

$$\|Jx\|^2 = \int_a^b |Jx(s)|^2 ds \leq \left(\int_a^b |\alpha + \beta t|^2 dt \right) \|x\|^2 \cdot \|w\|^2$$

donde l'assunto.

Se λ appartiene all'insieme $\{0, \alpha w_0 + \beta w_1\}$ l'operatore $J - \lambda I$ non è invertibile, dato che l'equazione integrale, come si è visto, quando ammette soluzioni ne ammette infinite. Pertanto i numeri complessi $0, \alpha w_0 + \beta w_1$ sono autovalori di J (naturalmente possono essere coincidenti).

Poichè gli autovalori costituiscono lo spettro puntuale, da quanto sopra consegue che lo spettro continuo e lo spettro residuo di J sono vuoti.

ESEMPIO 2.

Siano u, v elementi dello spazio $L^2(\Omega)$, linearmente indipendenti. Fissato $y \in L^2(\Omega)$, per ogni $\lambda \in \mathbb{C}$ consideriamo l'equazione integrale:

$$(21) \qquad \int_a^b [u(s) + t\,v(s)]\,x(t)\,dt - \lambda x(s) = y(s) \quad \text{q.o. in } \Omega.$$

L'operatore integrale J è generato dal nucleo

$$K(s,t) = u(s) + t\,v(s),$$

ed è definito ponendo per ogni $x \in L^2(\Omega)$:

$$(22) \qquad Jx(s) = \int_a^b [u(s) + t\,v(s)]\,x(t)\,dt \qquad \text{q.o. in } \Omega .$$

Useremo le notazioni y_0, y_1 introdotte con la seconda delle (14) [pag. 6], nonché le notazioni analoghe:

$$(23) \qquad \begin{cases} u_0 = \int_a^b u(t)\,dt \\ u_1 = \int_a^b t\,u(t)\,dt \end{cases} \qquad \begin{cases} v_0 = \int_a^b v(t)\,dt \\ v_1 = \int_a^b t\,v(t)\,dt \end{cases} .$$

Anche per lo studio della (21) esaminiamo prima il caso $\lambda \neq 0$.

Se $x \in L^2(\Omega)$ è un'eventuale soluzione della (21), posto:

$$(24) \qquad c_0 = \int_a^b x(t)\,dt \quad , \qquad c_1 = \int_a^b t\,x(t)\,dt$$

si ha:

$$c_0 u(s) + c_1 v(s) - \lambda x(s) = y(s)$$

e quindi:

$$(25) \qquad x = \frac{1}{\lambda}\left(c_0 u + c_1 v - y\right) .$$

Siccome x deve verificare la (21), ne segue:

$$\frac{1}{\lambda}\int_a^b [u(s) + t\,v(s)]\cdot[c_0 u(t) + c_1 v(t) - y(t)]\,dt - [c_0 u(s) + c_1 v(s) - y(s)] = y(s) \qquad \text{q.o. in } \Omega$$

ossia:

$$u(s)\,[c_0 u_0 + c_1 v_0 - \lambda c_0 - y_0] + v(s)\,[c_0 u_1 + c_1 v_1 - \lambda c_1 - y_1] = 0 \qquad \text{q.o. in } \Omega .$$

Essendo u e v linearmente indipendenti, i coefficienti tra parentesi quadre devono essere nulli, e quindi deve risultare:

$$(26) \qquad \begin{cases} (u_0 - \lambda)\,c_0 + v_0\,c_1 = y_0 \\ u_1\,c_0 + (v_1 - \lambda)\,c_1 = y_1 \end{cases} .$$

Dunque se x è un'eventuale soluzione della (21), necessariamente x ha l'espressione (25), con c_0 e c_1 soddisfacenti alle (26).

Viceversa con facili calcoli si prova che, se x è l'elemento di $L^2(\Omega)$ espresso dalla (25) con (c_0, c_1) soddisfacente alla (26), allora gli integrali a secondo membro delle (24) sono uguali rispettivamente a c_0 e c_1, e x è soluzione della (21).

Se λ non è autovalore della matrice:

$$(27) \qquad \begin{bmatrix} u_0 & v_0 \\ u_1 & v_1 \end{bmatrix}$$

esiste un'unica coppia (c_0, c_1) soddisfacente alle (26), e quindi l'equazione integrale (21) ammette una e una sola soluzione, data dalla (25) con (c_0, c_1) soluzione del sistema (26).

Se invece λ e' autovalore della matrice (27), e se le due matrici:

$$\begin{bmatrix} u_0-\lambda & v_0 \\ v_1 & v_1-\lambda \end{bmatrix}, \quad \begin{bmatrix} u_0-\lambda & v_0 & y_0 \\ v_1 & v_1-\lambda & y_1 \end{bmatrix}$$

hanno lo stesso rango, l'equazione integrale (21) ammette infinite soluzioni: infatti esistono infinite coppie (c_0, c_1) soddisfacenti alle (26), e quindi la (25) fornisce infiniti elementi di $L^2(\Omega)$ soluzioni della (21) (1).

Ancora per λ autovalore della matrice (27), se le due matrici anzidette non hanno lo stesso rango, non esiste alcuna coppia (c_0, c_1) soddisfacente alle (26) e quindi l'equazione integrale (21) non ammette soluzioni.

Passiamo ora allo studio della (21) con $\lambda = 0$, cioè dell'equazione integrale:

$$\int_a^b [u(s)+tv(s)]\, x(t)\,dt = y(s) \qquad \text{q.o. in } \Omega$$

che si può scrivere:

$$(28) \qquad u(s)\int_a^b x(t)\,dt + v(s)\int_a^b t\,x(t)\,dt = y(s) \qquad \text{q.o. in } \Omega.$$

Siccome il primo membro e' una combinazione lineare di u e v, ovviamente se il secondo membro y non e' combinazione lineare di u e v l'equazione (28) non ammette soluzioni.

Supponiamo allora che sia $y = h_0 u + h_1 v$, con $h_0, h_1 \in \mathbb{C}$. La (28) diviene:

$$u(s)\left[\int_a^b x(t)\,dt - h_0\right] + v(s)\left[\int_a^b t\,x(t)\,dt - h_1\right] = 0 \qquad \text{q.o. in } \Omega$$

e questa, per l'osservazione di pag. 5, ammette infinite soluzioni.

In definitiva per $\lambda = 0$ l'equazione (21) o non ammette soluzioni, o ne ammette infinite.

Teoria spettrale per l'operatore J (cfr. pag. 239)

Scriviamo l'equazione integrale (21) nella forma:

$$(J-\lambda I)x = y,$$

essendo J l'operatore integrale definito mediante la (22) ed I l'operatore identico in $L^2(\Omega)$.

Detti λ_1, λ_2 gli autovalori della matrice (27), ciascun numero complesso $\lambda \notin \{0, \lambda_1, \lambda_2\}$ appartiene all'insieme risolvente dell'operatore J. Avendo già visto che per un tale λ l'operatore $J-\lambda I$ e' invertibile ed ha per codominio tutto lo spazio $L^2(\Omega)$, si tratta di dimostrare che esiste una costante k tale che:

$$\|Jx\| \le k\|x\| \qquad \forall x \in L^2(\Omega)$$

in quanto ciò equivale alla continuità di J [(2.1) pag. 138], e comporta che λ appartiene

---

(1) E' immediato constatare che, se (c_0', c_1'), (c_0'', c_1'') sono soluzioni distinte del sistema (26), gli elementi di $L^2(\Omega)$:

$$x' = \frac{1}{\lambda}(c_0' u + c_1' v - y), \quad x'' = \frac{1}{\lambda}(c_0'' u + c_1'' v - y)$$

sono distinti. Infatti se fosse $x' = x''$, si avrebbe $(c_0' - c_0'')u + (c_1' - c_1'')v = 0$ e quindi $c_0' = c_0''$, $c_1' = c_1''$, il che e' assurdo.

all'insieme risolvente di J [n. 2 pag. 240].

Per la disuguaglianza di Schwarz si ha intanto:

$$\int_a^b |t|\cdot|x(t)|\,dt \le \left(\int_a^b t^2 dt\right)^{1/2} \|x\|$$

nonché:

$$\int_a^b |x(t)|\,dt = \int_a^b 1\cdot|x(t)|\,dt \le \left(\int_a^b 1^2 dt\right)^{1/2} \|x\| .$$

Dalla (22) consegue allora:

$$|Jx(s)| \le K_1 |u(s)|\cdot\|x\| + K_2 |v(s)|\cdot\|x\|$$

con ovvio significato di K_1 e K_2. Se ne deduce, tenendo presente la ovvia disuguaglianza $ab \le \frac{1}{2}(a^2+b^2)$, che risulta:

$$\|Jx\|^2 = \int_a^b |Jx(s)|^2 ds \le \frac{1}{2}K_1^2\|x\|^2\cdot\|u\|^2 + \frac{1}{2}K_2^2\|x\|^2\cdot\|v\|^2$$

da cui:

$$\|Jx\| \le \frac{1}{\sqrt{2}}\left(K_1\|u\| + K_2\|v\|\right)\|x\| .$$

Se λ appartiene all'insieme $\{0, \lambda_1, \lambda_2\}$ l'operatore $J-\lambda I$ non è invertibile, dato che l'equazione integrale, come si è visto, o non ammette soluzioni o ne ammette infinite.

Pertanto i numeri complessi $0, \lambda_1, \lambda_2$ sono autovalori di J (naturalmente due di essi, ed anche tutti, possono essere coincidenti).

Poiché gli autovalori costituiscono lo spettro puntuale, da quanto sopra consegue che lo spettro continuo e lo spettro residuo di J sono vuoti

ESEMPIO 3.

Per concludere questa "Introduzione" ritorniamo all'inizio del nostro discorso, riportando un esempio di "sistema lineare con infinite equazioni ed infinite incognite".

Consideriamo due successioni di numeri complessi

$$\sigma = (u_n)_{n\in N} \quad , \quad \tau = (v_n)_{n\in N}$$

elementi dello spazio ℓ_2 tali che per almeno una coppia $(p,q)\in N\times N$ si abbia:

$$\begin{vmatrix} u_p & u_q \\ v_p & v_q \end{vmatrix} \neq 0 \tag{29}$$

Per ogni $(i,j)\in N\times N$ denotiamo con a_{ij} il prodotto scalare in $\mathbb{C}\times\mathbb{C}$ delle coppie (u_i, v_i), (u_j, v_j), cioè poniamo:

$$a_{ij} = u_i\bar{u}_j + v_i\bar{v}_j .$$

Verifichiamo che per la matrice $M = \{a_{ij}\}$ il limite di pag. 1 denotato con $\|M\|$, ovviamente non nullo, è finito.

Dalla relazione:

$$|a_{ij}|^2 \le \left(|u_i|\cdot|u_j| + |v_i|\cdot|v_j|\right)^2 = |u_i|^2|u_j|^2 + |v_i|^2|v_j|^2 + 2|u_i|\cdot|v_i|\cdot|u_j|\cdot|v_j| ,$$

tenendo presente che per la disuguaglianza di Schwarz in $\mathbb{R}^m$ risulta

$$\sum_{j=1}^{m} |u_j|\cdot|v_j| \le \left(\sum_{j=1}^{m}|u_j|^2\right)^{1/2}\left(\sum_{j=1}^{m}|v_j|^2\right)^{1/2} \le \|\sigma\|\cdot\|\tau\|,$$

si deduce:

$$\sum_{j=1}^{m}|a_{ij}|^2 \le |u_i|^2\cdot\|\sigma\|^2+|v_i|^2\cdot\|\tau\|^2+2|u_i|\cdot|v_i|\sum_{j=1}^{m}|u_j|\cdot|v_j| \le |u_i|^2\cdot\|\sigma\|^2+|v_i|^2\cdot\|\tau\|^2+2|u_i|\cdot|v_i|\cdot\|\sigma\|\cdot\|\tau\|.$$

Ne segue:

$$\sum_{i,j}^{1\dots m}|a_{ij}|^2 \le \|\sigma\|^2\sum_{i=1}^{m}|u_i|^2+\|\tau\|^2\sum_{i=1}^{m}|v_i|^2+2\|\sigma\|\cdot\|\tau\|\cdot\sum_{i=1}^{m}|u_i|\cdot|v_i| \le$$

$$\le \|\sigma\|^4+\|\tau\|^4+2\|\sigma\|^2\cdot\|\tau\|^2=\left(\|\sigma\|^2+\|\tau\|^2\right)^2$$

donde l'asserto.

Dopo ciò, assegnata la successione $(y_n)_{n\in\mathbb{N}} \in \ell_2$, per ogni $\lambda\in\mathbb{C}$ ha senso considerare il sistema nell'incognita $(x_n)_{n\in\mathbb{N}}$ [si ricordi quanto stabilito a pag. 1]:

$$(30)\qquad \begin{cases} (a_{11}-\lambda)x_1 + a_{12}x_2 + \dots + a_{1n}x_n + \dots = y_1 \\ a_{21}x_1 + (a_{22}-\lambda)x_2 + \dots + a_{2n}x_n + \dots = y_2 \\ \dots\dots\dots\dots\dots\dots\dots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + (a_{mn}-\lambda)x_n + \dots = y_m \\ \dots\dots\dots\dots\dots\dots\dots \end{cases}$$

Al fine di studiare il sistema (30), posto:

$$\Omega = \bigcup_{i\in\mathbb{N}}]i, i+1[,$$

consideriamo l'operatore $B : \ell_2 \to L^2(\Omega)$ introdotto a pag. 2, e denotiamo con u, v, x, y rispettivamente le immagini mediante B delle successioni $(u_n)_{n\in\mathbb{N}}$, $(v_n)_{n\in\mathbb{N}}$, $(x_n)_{n\in\mathbb{N}}$, $(y_n)_{n\in\mathbb{N}}$, sicché u, v, x, y hanno un rappresentante che in ciascun intervallo $]i, i+1[$ è costante, ed è definito ponendo:

$$(31)\qquad u(s)=u_i \;,\quad v(s)=v_i \;,\quad x(s)=x_i \;,\quad y(s)=y_i \qquad \forall s\in\,]i,i+1[.$$

È importante evidenziare che per l'ipotesi (29) i vettori u, v di $L^2(\Omega)$ sono linearmente indipendenti.

Tenuto conto di quanto osservato nella nota (1) di pag. 2, per le (31) il primo membro dell'equazione n^{ma} del sistema (30) diviene:

$$\sum_{j=1}^{\infty} a_{nj}x_j - \lambda x_n = \sum_{j=1}^{\infty}(u_n\bar{u}_j+v_n\bar{v}_j)x_j - \lambda x_n = u_n\sum_{j=1}^{\infty}x_j\bar{u}_j + v_n\sum_{j=1}^{\infty}x_j\bar{v}_j - \lambda x_n = (x,u)u_n+(x,v)v_n-\lambda x_n$$

e quindi il sistema (30) si può scrivere:

$$(x,u)u_n+(x,v)v_n-\lambda x_n = y_n \qquad \forall n\in\mathbb{N} .$$

Pertanto il sistema è equivalente all'equazione nell'incognita $x\in L^2(\Omega)$:

$$(32)\qquad (x,u)u+(x,v)v-\lambda x = y .$$

Si tratta dell'equazione integrale:

$$\int_{\Omega}\left[u(s)\bar{u}(t)+v(s)\bar{v}(t)\right]x(t)\,dt-\lambda x(s)=y(s) \qquad \text{q.o. in } \Omega ,$$

dove l'operatore integrale J è generato dal nucleo

$$K(s,t) = u(s)\bar{u}(t) + v(s)\bar{v}(t).$$

Per lo studio della (32) esaminiamo dapprima il caso $\lambda \neq 0$.

Se x è un'eventuale soluzione della (32), posto:

(33) $$c_1 = (x,u) \quad , \quad c_2 = (x,v)$$

si ha:

$$c_1 u + c_2 v - \lambda x = y$$

e quindi:

(34) $$x = \frac{1}{\lambda}(c_1 u + c_2 v - y)$$

Siccome x deve verificare la (32), sostituendo (34) in (32) si perviene all'uguaglianza:

$$u\left[(\|u\|^2 - \lambda)c_1 + (v,u)c_2 - (y,u)\right] + v\left[(u,v)c_1 + (\|v\|^2 - \lambda)c_2 - (y,v)\right] = 0$$

dalla quale, essendo u, v linearmente indipendenti, consegue:

$$(\|u\|^2 - \lambda)c_1 + (v,u)c_2 - (y,u) = 0 \quad , \quad (u,v)c_1 + (\|v\|^2 - \lambda)c_2 - (y,v) = 0$$

ossia:

(35) $$\begin{cases} (\|u\|^2 - \lambda)c_1 + (v,u)c_2 = (y,u) \\ (u,v)c_1 + (\|v\|^2 - \lambda)c_2 = (y,v) \end{cases}.$$

Dunque se x è un'eventuale soluzione della (32), necessariamente x ha l'espressione (34), con c_1, c_2 soddisfacenti al sistema (35).

Viceversa con facili calcoli si prova che, se x è l'elemento di $L^2(\Omega)$ espresso dalla (34) con (c_1, c_2) soluzione del sistema (35), allora i prodotti scalari a secondo membro della (33) sono uguali rispettivamente a c_1 e c_2, e x è soluzione della (32).

Se λ non è autovalore della matrice:

(36) $$\begin{bmatrix} \|u\|^2 & (v,u) \\ (u,v) & \|v\|^2 \end{bmatrix}$$

esiste un'unica coppia (c_1, c_2) soluzione del sistema (35), e quindi l'equazione integrale (32) ammette una e una sola soluzione, data dalla (34) con (c_1, c_2) soluzione del sistema (35).

Se invece λ è autovalore della matrice (36), e se le due matrici:

$$\begin{bmatrix} \|u\|^2 - \lambda & (v,u) \\ (u,v) & \|v\|^2 - \lambda \end{bmatrix} \quad , \quad \begin{bmatrix} \|u\|^2 - \lambda & (v,u) & (y,u) \\ (u,v) & \|v\|^2 - \lambda & (y,v) \end{bmatrix}$$

hanno lo stesso rango, l'equazione (32) ammette infinite soluzioni: infatti esistono infinite coppie (c_1, c_2) soluzioni del sistema (35), e quindi la (34) fornisce infiniti elementi di $L^2(\Omega)$ soluzioni della (32) (1).

Ancora per λ autovalore della matrice (36), se le due matrici anzidette non hanno lo stesso rango, non esiste alcuna soluzione del sistema (35) e quindi l'equazione (32) non ammette soluzioni.

---

(1) Se (c_1', c_2'), (c_1'', c_2'') sono soluzioni distinte del sistema (35), la (34) fornisce soluzioni distinte della (32): ciò, essendo u e v linearmente indipendenti, si prova come nella nota (1) di pag. 10.

* * *

Ritorniamo ora, sempre nel caso $\lambda \neq 0$, al sistema lineare (30) nell'incognita $(x_n)_{n\in\mathbb{N}}$ che equivale, come si è visto a pag. 12, all'equazione (32) nell'incognita x [con l'intesa che sussistano le (31)].

Se λ non è autovalore della matrice (36), siccome le funzioni u, v, y sono costanti a tratti, definite in ciascun intervallo $]i, i+1[$ dalla (31), per la (34) la soluzione x della (32) è a sua volta costante a tratti. Posto allora:

$$x(s) = x_n \qquad \forall s \in]n, n+1[\tag{37}$$

dalla (34) si deduce:

$$x_n = \frac{1}{\lambda}(c_1 u_n + c_2 v_n - y_n) \tag{38}$$

e questa fornisce la soluzione del sistema (30). Ricordiamo che nella (34), e quindi nella (38), la coppia (c_1, c_2) è la soluzione del sistema (35).

Se λ è autovalore della matrice (36), e se le due matrici dianzi considerate hanno lo stesso rango, per ciascuna soluzione (c_1, c_2) del sistema (35) la (38) fornisce una soluzione del sistema (30)

* * *

Passando allo studio della (32) con $\lambda = 0$, cioè dell'equazione integrale:

$$(x,u)u + (x,v)v = y\,, \tag{39}$$

osserviamo intanto che la (39) non ammette soluzioni se il termine noto y non è una combinazione lineare di u e v (in quanto il primo membro è una combinazione lineare di u e v). Perciò supponiamo che il termine noto y sia:

$$y = h_1 u + h_2 v \tag{40}$$

con $h_1, h_2 \in \mathbb{C}$, sicché la (39) diviene:

$$u[(x,u) - h_1] + v[(x,v) - h_2] = 0\,.$$

Dunque se x è un'eventuale soluzione della (39), essendo u e v linearmente indipendenti necessariamente x è soluzione del sistema:

$$\begin{cases} (x,u) = h_1 \\ (x,v) = h_2 \end{cases} \tag{41}$$

Viceversa, se x è soluzione del sistema (41), banalmente x verifica la (39), stante la (40).

Osserviamo che, mentre per $\lambda \neq 0$ le soluzioni della (32) sono necessariamente costanti a tratti [sussistendo la (37)], le eventuali soluzioni per $\lambda = 0$, cioè le eventuali soluzioni del sistema (41), non sono necessariamente costanti a tratti. Peraltro, come vedremo, esistono infinite soluzioni costanti a tratti del sistema (41).

Ricerchiamo dunque una successione $(x_n)_{n\in N}$ tale che, posto:

$$x(\vartheta) = x_n \quad \forall \vartheta \in]n, n+1[\ ,$$

la funzione $x \in L^2(\Omega)$ verifichi il sistema (41).

Siccome i prodotti scalari in $L^2(\Omega)$ che figurano nella (41) sono uguali ai prodotti scalari in ℓ_2 della successione $(x_n)_{n\in N}$ rispettivamente per le successioni $(u_n)_{n\in N}$, $(v_n)_{n\in N}$ [cfr. nota (1) pag. 2], il sistema (41) si può scrivere:

$$(42) \qquad \sum_{n\in N} x_n \bar{u}_n = h_1 \ , \quad \sum_{n\in N} x_n \bar{v}_n = h_2 \ .$$

Ricordando l'ipotesi (29) di pag. 11 si vede subito che tale sistema ammette infinite soluzioni. Infatti, posto $N' = N - \{p, q\}$ e scelto arbitrariamente x_n per ogni $n \in N'$, il sistema (42) si riduce al sistema nell'incognita (x_p, x_q):

$$\begin{cases} \bar{u}_p x_p + \bar{u}_q x_q = h_1 - \sum_{n\in N'} x_n \bar{u}_n \\ \bar{v}_p x_p + \bar{v}_q x_q = h_2 - \sum_{n\in N'} x_n \bar{v}_n \end{cases}$$

che ammette un'unica soluzione per la predetta ipotesi (29).

Pertanto il sistema (41) ammette infinite soluzioni costanti a tratti, e lo stesso accade per la (39) equivalente, come si è visto, al sistema (41).

Dopo ciò è chiaro che <u>per $\lambda = 0$ il sistema lineare (30) ammette infinite soluzioni (se e solo se y ha l'espressione (40), altrimenti non ha soluzioni)</u>.

<u>OSSERVAZIONE</u>

<u>Se le assegnate successioni $(u_n)_{n\in N}$ e $(v_n)_{n\in N}$, vettori dello spazio ℓ_2, sono ortogonali e di norma unitaria</u>, per quanto osservato nella nota (1) di pag. 2 anche le funzioni u e v definite mediante le (31) [pag. 12] sono, nello spazio $L^2(\Omega)$, vettori ortogonali e di norma unitaria. Di conseguenza la matrice (36):

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

ammette l'unico autovalore $\lambda = 1$ (di molteplicità 2), e il sistema (35) diviene:

$$(43) \qquad (1-\lambda) c_1 = (y, u) \ , \quad (1-\lambda) c_2 = (y, v) \ .$$

Quindi per $\lambda \notin \{0, 1\}$ la soluzione del sistema (30) è data da [cfr. (38)]:

$$(44) \qquad x_n = \frac{1}{\lambda(1-\lambda)} \left[(y, u) u_n + (y, v) v_n \right] - \frac{1}{\lambda} y_n \quad \forall n \in N \ .$$

Osserviamo inoltre che per $\lambda = 1$ la condizione di risolubilità del sistema (30) [che le due matrici di pag. 13 abbiano lo stesso rango] diviene:

$$(45) \qquad (y, u) = 0 \ , \quad (y, v) = 0 \ .$$

In tale ipotesi qualunque coppia (c_1, c_2) di numeri complessi verifica le (43), e quindi il sistema (30) ammette le infinite soluzioni date dalla formula [cfr. (38)]:

$$x_n = c_1 u_n + c_2 v_n - y_n \quad \forall n \in N$$

con c_1, c_2 costanti arbitrarie.

Per dare un esempio sulle precedenti osservazioni, per ogni $n \in \mathbb{N}$ poniamo:

$$\begin{cases} \text{se } n \text{ è dispari}: & u_n = a^{n+1}, \quad v_n = 0, & y_n = a^{n+1} \\ \text{se } n \text{ è pari}: & u_n = 0, \quad v_n = a^n, & y_n = a^n \end{cases}$$

con

$$a = 1/\sqrt[4]{2} .$$

Poiché risulta:

$$\sum_{n=1}^{\infty} u_n^2 = \sum_{h=1}^{\infty} u_{2h-1}^2 = \sum_{h=1}^{\infty} (a^{2h})^2 = a^4 \sum_{h=1}^{\infty} (a^4)^{h-1} = \frac{a^4}{1-a^4} = 1$$

$$\sum_{n=1}^{\infty} v_n^2 = \sum_{h=1}^{\infty} v_{2h}^2 = \sum_{h=1}^{\infty} a^{4h} = 1$$

$$\sum_{n=1}^{\infty} y_n^2 = \sum_{h=1}^{\infty} y_{2h-1}^2 + \sum_{h=1}^{\infty} y_{2h}^2 = \sum_{h=1}^{\infty} a^{4h} + \sum_{h=1}^{\infty} a^{4h} = 2 ,$$

le successioni $(u_n)_{n\in\mathbb{N}}$, $(v_n)_{n\in\mathbb{N}}$, $(y_n)_{n\in\mathbb{N}}$ sono vettori dello spazio ℓ_2, i primi due ortogonali e di norma unitaria.

Per ogni $n \in \mathbb{N}$ il coefficiente di x_n nell'equazione k^{ma} del sistema lineare (30), con $n \neq k$:

$$a_{kn} = u_k u_n + v_k v_n$$

è nullo se k è dispari per ogni n pari (dato che $u_n = 0$ e $v_k = 0$), e se k è pari per ogni n dispari (dato che $u_k = 0$ e $v_n = 0$).

Essendo [1]:

$$(y,u) = \sum_{n\in\mathbb{N}} y_n u_n = \sum_{h=1}^{\infty} y_{2h-1} u_{2h-1} = \sum_{h=1}^{\infty} a^{2h} \cdot a^{2h} = 1$$

$$(y,v) = \sum_{n\in\mathbb{N}} y_n v_n = \sum_{h=1}^{\infty} y_{2h} u_{2h} = \sum_{h=1}^{\infty} a^{4h} = 1$$

per $\lambda \notin \{0,1\}$ la soluzione del sistema (30) è data da [cfr. (44)]:

$$x_n = \begin{cases} a^{n+1}/(1-\lambda) & \text{se } n \text{ è dispari} \\ a^n/(1-\lambda) & \text{se } n \text{ è pari} \end{cases} \qquad \forall n \in \mathbb{N} .$$

Per $\lambda = 1$ il sistema (30) non ammette soluzioni, perché non è verificata la condizione (45). Anche per $\lambda = 0$ non ammette soluzioni, perché l'elemento y di $L^2(\Omega)$ individuato dai termini noti non è del tipo (40).

---

(1) Si ricordi che y, u, v sono costanti in ciascun intervallo $]i, i+1[$, e ivi definiti dalle (31) di pag. 12. Si tenga presente altresì quanto osservato nella nota (1) di pag. 2.

OPERATORI INTEGRALI
EQUAZIONI INTEGRALI DI FREDHOLM.

1. Siano $\Omega_1 \subset R^k$, $\Omega_2 \subset R^h$ aperti limitati, e $K(s,t)$ un elemento di $L^2(\Omega_1 \times \Omega_2)$. A pag. 169 abbiamo introdotto gli operatori integrali J_1 e J_2 generati dal nucleo $K(s,t)$, il primo chiamato "con variabile di integrazione t", il secondo "con variabile di integrazione s". Poichè lo studio di ciascuno di essi è del tutto analogo a quello dell'altro, qui ci riferiremo soltanto all'operatore J_1, che indicheremo semplicemente con J, e riterremo sottintesa la precisazione "con variabile di integrazione t". Tenendo conto dei nostri scopi, supporremo $\Omega_1 = \Omega_2 = \Omega$.

In primo luogo richiamiamo la definizione dell'operatore J.

Per ogni $x \in L^2(\Omega)$, il corrispondente Jx è la funzione definita q.o. in Ω ponendo:

$$(1) \qquad Jx(s) = \int_\Omega K(s,t)x(t)dt \qquad \text{q.o. in } \Omega.$$

A pag. 169 abbiamo visto che per ogni $x \in L^2(\Omega)$ anche Jx appartiene ad $L^2(\Omega)$, sicchè J è un operatore di $L^2(\Omega)$ in $L^2(\Omega)$, ovviamente lineare. Abbiamo stabilito che si ha:

$$(2) \qquad \|Jx\|_{L^2(\Omega)} \le \|K\|_{L^2(\Omega\times\Omega)} \cdot \|x\|_{L^2(\Omega)} \qquad \forall x \in L^2(\Omega)$$

e quindi che J è anche continuo. Pertanto J è un elemento dello spazio $\mathcal{L}(L^2(\Omega), L^2(\Omega))$ [pag. 203], e per la (2) risulta:

$$(3) \qquad \|J\|_{\mathcal{L}(L^2(\Omega),L^2(\Omega))} \le \|K\|_{L^2(\Omega\times\Omega)}.$$

Abbiamo stabilito che J è un operatore compatto [pag. 201]. L'aggiunto di J [pag. 220], come si è visto nell'es. 5 di pag. 227.4, è l'operatore integrale generato dal nucleo $\overline{K(t,s)}$ (trasposto del coniugato), sicchè per ogni $x \in L^2(\Omega)$ si ha:

$$J^*x(s) = \int_\Omega \overline{K(t,s)}\, x(t)dt \qquad \text{q.o. in } \Omega.$$

Se risulta $\overline{K(t,s)} = K(s,t)$ (nucleo simmetrico, o hermitiano), l'operatore integrale J è autoaggiunto.

D'ora in avanti le norme in $L^2(\Omega)$, in $L^2(\Omega\times\Omega)$ e in $\mathcal{L}(L^2(\Omega), L^2(\Omega))$ saranno indicate tutte con $\|\cdot\|$, non essendovi possibilità di equivoci. Peraltro in qualche caso sarà opportuno, per motivi di chiarezza, evidenziare lo spazio al quale la norma si riferisce.

Denoteremo con I l'operatore identico di $L^2(\Omega)$ in $L^2(\Omega)$. Fissati $\lambda \in \mathbb{C}$ ed $y \in L^2(\Omega)$, l'equazione nell'incognita $x \in L^2(\Omega)$:

$$(J-\lambda I)x = y \tag{4}$$

si chiama <u>equazione integrale di Fredholm</u>. Si tratta del problema di stabilire se esiste almeno un $x \in L^2(\Omega)$ tale che:

$$\int_\Omega K(s,t)x(t)\,dt - \lambda x(s) = y(s) \qquad \text{q.o. in } \Omega. \tag{5}$$

Lo studio della (4) si fonda su quello degli operatori J e $J-\lambda I$, sicché le considerazioni seguenti riguardano prevalentemente tali operatori.

2. Vogliamo mostrare che per ogni $n \in N$ la potenza J^n dell'operatore J è a sua volta un operatore integrale, generato da un opportuno nucleo.

In via preliminare definiamo induttivamente una successione $\{K_n\}$ di elementi di $L^2(\Omega\times\Omega)$, chiamati <u>i nuclei iterati di K</u>. Posto:

$$K_0(s,t) = K(s,t) \quad \text{q.o. in } \Omega\times\Omega,$$

per $n=1$ è vera la proposizione:

(6) <u>E' stato definito il nucleo $K_{n-1}(s,t)$ e questo appartiene ad $L^2(\Omega\times\Omega)$</u>,

e quindi ha significato porre, per $n=1$:

$$K_n(s,t) = \int_\Omega K(s,\xi)K_{n-1}(\xi,t)\,d\xi \quad \text{q.o. in } \Omega\times\Omega. \tag{7}$$

Procedendo per induzione, proveremo poi che per ogni $n \in N$ ha significato definire K_n mediante la (7), che $K_n \in L^2(\Omega\times\Omega)$ ed inoltre che risulta:

$$\|K_n\| \le \|K\|^{n+1} \qquad \forall n \in N. \tag{8}$$

DIM. Tenendo presente la disuguaglianza di Schwarz, si ha:

$$|K_1(s,t)|^2 = \left|\int_\Omega K(s,\xi)K_0(\xi,t)\,d\xi\right|^2 \le \left(\int_\Omega |K(s,\xi)|\cdot|K_0(\xi,t)|\,d\xi\right)^2 \le \left(\int_\Omega |K(s,\xi)|^2 d\xi\right)\left(\int_\Omega |K_0(\xi,t)|^2 d\xi\right)$$

e quindi:

$$\iint_{\Omega\times\Omega} |K_1(s,t)|^2\,ds\,dt \le \iint_{\Omega\times\Omega} \underbrace{\left(\int_\Omega |K(s,\xi)|^2 d\xi\right)}_{f(s)}\underbrace{\left(\int_\Omega |K_0(\xi,t)|^2 d\xi\right)}_{g(t)} ds\,dt = \iint_{\Omega\times\Omega} f(s)g(t)\,ds\,dt$$

$$= \int_\Omega f(s)\,ds \int_\Omega g(t)\,dt = \left(\int_\Omega f(s)\,ds\right)\left(\int_\Omega g(t)\,dt\right) = \iint_{\Omega\times\Omega} |K(s,\xi)|^2 d\xi\,ds \cdot \iint_{\Omega\times\Omega} |K_0(\xi,t)|^2 d\xi\,dt = \|K\|^2\cdot\|K_0\|^2 = \|K\|^4.$$

Ciò prova che $K_1 \in L^2(\Omega\times\Omega)$, e che è vera la (8) per $n=1$.

Procedendo per induzione, ammettiamo che per $n=m$ $(m>1)$ sia vera la (6), sicché ha senso la (7), ed inoltre che sussista la (8) con $n=m-1$. Sostituendo formalmente nelle precedenti disuguaglianze K_0 con K_{m-1} e K_1 con K_m, se ne deduce che $K_m \in L^2(\Omega\times\Omega)$ e che la (8) è vera anche per $n=m$. Dopo ciò l'asserto è provato.

Orbene, dimostriamo che:

(2.1) Per ogni $r \in N$ l'operatore J^r è l'operatore integrale generato dal nucleo $K_{r-1}(s,t)$, cioè per $x \in L^2(\Omega)$ si ha, quasi ovunque in Ω [cfr. (1)]:

$$(9) \qquad J^r_x(s) = \int_\Omega K_{r-1}(s,t)\,x(t)\,dt \qquad \forall r \in N.$$

DIM. Ricordando la (1) e l'uguaglianza $K_0 = K$, banalmente la (9) è vera per $r=1$. Procedendo per induzione, ammettiamo che J^{i-1} sia l'operatore integrale generato dal nucleo K_{i-2} ($i \geq 2$), cioè che per $x \in L^2(\Omega)$ sia vera la (9) con $r = i-1$:

$$(10) \qquad J^{i-1}_x(\xi) = \int_\Omega K_{i-2}(\xi,t)\,x(t)\,dt \qquad \text{q.o. in } \Omega.$$

Per comodità, detto z un qualunque elemento di $L^2(\Omega)$, scriviamo la (1) con z in luogo di x:

$$(11) \qquad Jz(s) = \int_\Omega K(s,\xi)\,z(\xi)\,d\xi \qquad \text{q.o. in } \Omega.$$

Poiché J^i_x significa $J(J^{i-1}_x)$, per la (11) con $z = J^{i-1}_x$ si ha:

$$J^i_x(s) = J(J^{i-1}_x)(s) = \int_\Omega K(s,\xi)\left[J^{i-1}_x(\xi)\right]d\xi$$

e quindi, per la (10):

$$J^i_x(s) = \int_\Omega K(s,\xi)\,d\xi \int_\Omega K_{i-2}(\xi,t)\,x(t)\,dt.$$

Ne segue che:

$$J^i_x(s) = \iint_{\Omega\times\Omega} K(s,\xi)K_{i-2}(\xi,t)\,x(t)\,dt\,d\xi = \int_\Omega x(t)\,dt \int_\Omega K(s,\xi)K_{i-2}(\xi,t)\,d\xi$$

Per la (7) l'integrale interno è uguale a $K_{i-1}(s,t)$, e perciò l'ultima uguaglianza diviene:

$$J^i_x(s) = \int_\Omega K_{i-1}(s,t)\,x(t)\,dt \qquad \text{q.o. in } \Omega.$$

Dunque la (9) è vera anche per $r = i$, e con ciò l'assunto è provato.

Se $|\lambda| > \|K\|$, le serie:

$$(12) \qquad \sum_{r=1}^{\infty} \frac{J^r}{\lambda^r} \quad , \quad \sum_{r=1}^{\infty} \frac{K_{r-1}}{\lambda^r}$$

sono entrambe convergenti (ciascuna nel suo spazio).

DIM. Per la (12) di pag. 108 e poi per la (3) si ha:

$$\|J^r/\lambda^r\| \leq \|J\|^r/|\lambda|^r \leq (\|K\|/|\lambda|)^r$$

e d'altra parte per la (8) risulta:

$$\|K_{r-1}/\lambda^r\| \leq \|K\|^r/|\lambda|^r = (\|K\|/|\lambda|)^r.$$

Dunque per le (12) accade che la serie delle norme è maggiorata da una serie convergente [dato che $(\|K\|/|\lambda|) < 1$].

Ne segue per il teor. (3.1) di pag. 215 la convergenza della prima serie, e per il teor. (6.1) di pag. 125 la convergenza della seconda. La condizione menzionata in (6.1) [condizione di Cauchy per le serie] è soddisfatta in quanto verificata dalla serie maggiorante convergente.

Per $|\lambda|>\|K\|$ poniamo:

$$J^{(\lambda)} = \sum_{\nu=1}^{\infty} J^\nu/\lambda^\nu \quad , \quad K^{(\lambda)} = \sum_{\nu=1}^{\infty} K_{\nu-1}/\lambda^\nu . \tag{13}$$

La prima delle uguaglianze (13) va intesa nel senso dello spazio $\mathcal{L}(L^2(\Omega),L^2(\Omega))$; essa sussiste anche puntualmente in $L^2(\Omega)$ [si veda pag.217], cioè si ha:

$$J^{(\lambda)}y = \sum_{\nu=1}^{\infty} J^\nu y/\lambda^\nu \qquad \forall y\in L^2(\Omega) \tag{14}$$

e questa va intesa nel senso di $L^2(\Omega)$ [i termini della serie sono punti di $L^2(\Omega)$].

La seconda delle (13) prende il nome di serie di Neumann relativa al nucleo K, e naturalmente va intesa nel senso di $L^2(\Omega\times\Omega)$.

Fissato $\lambda\in\mathbb{C}-\{0\}$, dalla (2.1) consegue che per ogni $\nu\in\mathbb{N}$ i termini di indice ν delle serie (13) sono legati dalla proprietà:

J^ν/λ^ν è l'operatore integrale generato dal nucleo $K_{\nu-1}/\lambda^\nu$.

Orbene, per $|\lambda|>\|K\|$ anche le somme delle due serie sono legate dalla stessa proprietà. Ciò sussiste il teorema.

(2.2). Se $|\lambda|>\|K\|$, l'operatore $J^{(\lambda)}$ è un operatore integrale, generato dal nucleo $K^{(\lambda)}$ somma della serie di Neumann. In altri termini, qualunque sia $y\in L^2(\Omega)$ si ha:

$$J^{(\lambda)}y(s) = \int_\Omega K^{(\lambda)}(s,t)\,y(t)\,dt \quad \text{q.o. in } \Omega . \tag{15}$$

DIM. Detto $F^{(\lambda)}$ l'operatore integrale generato dal nucleo $K^{(\lambda)}$, per ogni $y\in L^2(\Omega)$ il secondo membro della (15) è $F^{(\lambda)}y(s)$. L'uguaglianza $J^{(\lambda)}=F^{(\lambda)}$ significa che per ogni $y\in L^2(\Omega)$ risulta $J^{(\lambda)}y=F^{(\lambda)}y$, ossia per la (14):

$$\lim_{n\to\infty} \sum_{\nu=1}^{n} J^\nu y/\lambda^\nu = F^{(\lambda)}y \quad \text{in } L^2(\Omega) . \tag{16}$$

Dunque, esplicitando la (16), si tratta di dimostrare che:

$$\lim_{n\to\infty} \int_\Omega \left| \left(\sum_{\nu=1}^{n} J^\nu y/\lambda^\nu\right)(s) - \int_\Omega K^{(\lambda)}(s,t)\,y(t)\,dt \right|^2 ds = 0 . \tag{17}$$

Ci occorrerà esplicitare la seconda delle (13). Detta σ_n la somma parziale n^{ma} della serie, cioè posto:

$$\sigma_n = \sum_{\nu=1}^{n} K_{\nu-1}/\lambda^\nu ,$$

la seconda delle (13) significa che:

$$\lim_{n\to\infty} \|\sigma_n - K^{(\lambda)}\|_{L^2(\Omega\times\Omega)} = 0 . \tag{18}$$

Ciò premesso, valutiamo l'integrando a primo membro della (17). Tenuto conto

della (9), e successivamente utilizzando la disuguaglianza di Schwarz, si ha:

$$\left|\left(\sum_{r=1}^{n}\frac{J_r^2}{\lambda^2}\right)(s)-\int_\Omega K^{(\lambda)}(s,t)y(t)\,dt\right|^2=\left|\sum_{r=1}^{n}\int_\Omega\frac{K_{r-1}(s,t)}{\lambda^r}y(t)\,dt-\int_\Omega K^{(\lambda)}(s,t)y(t)\,dt\right|^2=$$

$$=\left|\int_\Omega\sigma_n(s,t)y(t)\,dt-\int_\Omega K^{(\lambda)}(s,t)y(t)\,dt\right|^2=\left|\int_\Omega[\sigma_n(s,t)-K^{(\lambda)}(s,t)]y(t)\,dt\right|^2=$$

$$=\left|\left(\sigma_n(s,\cdot)-K^{(\lambda)}(s,\cdot),\overline{y}\right)_{L^2(\Omega)}\right|^2\le\|\sigma_n(s,\cdot)-K^{(\lambda)}(s,\cdot)\|^2_{L^2(\Omega)}\cdot\|y\|^2_{L^2(\Omega)}.$$

Integrando su Ω, si ottiene:

$$\int_\Omega\left|\left(\sum_{r=1}^{n}\frac{J_r^2}{\lambda^2}\right)(s)-\int_\Omega K^{(\lambda)}(s,t)y(t)\,dt\right|^2ds\le\|y\|^2_{L^2(\Omega)}\int_\Omega\|\sigma_n(s,\cdot)-K^{(\lambda)}(s,\cdot)\|^2_{L^2(\Omega)}\,ds$$

$$=\|y\|^2_{L^2(\Omega)}\int_\Omega ds\int_\Omega|\sigma_n(s,t)-K^{(\lambda)}(s,t)|^2dt=\|y\|^2_{L^2(\Omega)}\iint_{\Omega\times\Omega}|\sigma_n(s,t)-K^{(\lambda)}(s,t)|^2ds\,dt=$$

$$=\|y\|_{L^2(\Omega)}\cdot\|\sigma_n-K^{(\lambda)}\|^2_{L^2(\Omega\times\Omega)}.$$

Siccome per $n\to\infty$ l'ultimo membro tende a 0 per la (18), dalla precedente relazione si deduce la (17) e l'assunto è dimostrato.

La (15) fornisce un'espressione di $J_y^{(\lambda)}(s)$: da notare che un'espressione di tale funzione non si può ottenere dalla (14), perché non si sa che la serie a secondo membro, di punti di $L^2(\Omega)$, converge quasi ovunque in Ω. Un noto teorema asserisce che, <u>se una successione $\{u_n\}$ di punti di $L^2(\Omega)$ converge (in $L^2(\Omega)$) ad u_0, si può estrarre dalla $\{u_n\}$ una successione convergente ad u_0 quasi ovunque in Ω</u>: ma tale teorema, applicato alla successione delle somme parziali della (14), ovviamente non basta per ottenere un'espressione di $J_y^{(\lambda)}(s)$.

Dalle considerazioni precedenti emerge l'importanza del fatto che l'operatore $J^{(\lambda)}$, somma di una serie, è generato dal nucleo $K^{(\lambda)}$.

3. Denotiamo con I l'operatore identico di $L^2(\Omega)$ in $L^2(\Omega)$. Se $\lambda\neq0$, essendo J compatto, per l'equazione integrale $(J-\lambda I)x=y$ vale la tesi del teorema dell'alternativa [(6.1) pag. 233], tenuto conto di quanto detto a pag. 235 (secondo comma). Il verificarsi della prima eventualità menzionata nel teorema comporta che l'operatore $J-\lambda I$

è invertibile ed ha per codominio tutto lo spazio $L^2(\Omega)$, cioè che λ appartiene all'insieme risolvente dell'operatore J [n.2 pag.240]; il verificarsi della seconda eventualità comporta che λ è autovalore di J, ossia appartiene allo spettro puntuale di J (pag.240).

Pertanto, con riferimento all'operatore integrale J, ciascun numero complesso $\lambda \neq 0$ che appartenga allo spettro è un autovalore [ed ha molteplicità finita, per la (3.2) di pag.244]; invece 0, punto dello spettro [inizio pag.242] può darsi che sia autovalore [J non invertibile], ma può anche accadere che appartenga allo spettro residuo [J invertibile, codominio di J non denso in $L^2(\Omega)$] oppure allo spettro continuo [J invertibile, codominio di J denso in $L^2(\Omega)$ (quindi J^{-1} non continuo, altrimenti 0 apparterrebbe all'insieme risolvente)].

Brevemente, detto Λ l'insieme degli autovalori di J, lo spettro di J è uguale a $\Lambda \cup \{0\}$. Per la citata (3.2) di pag.244, l'insieme Λ è finito o numerabile, e nel secondo caso ha come unico punto di accumulazione 0.

Il teorema che segue comporta che ogni numero complesso di modulo maggiore di $\|K\|$ appartiene all'insieme risolvente di J, sicché il raggio spettrale di J è minore o uguale a $\|K\|$, e risulta $|\lambda| \leq \|K\| \ \forall \lambda \in \Lambda$.

(3.1). Nell'ipotesi $|\lambda| > \|K\|$, l'operatore $J - \lambda I$ è invertibile ed ha per codominio $L^2(\Omega)$. L'operatore inverso si esprime come segue:

$$(19) \qquad (J-\lambda I)^{-1} y = -\frac{y}{\lambda} - \frac{1}{\lambda} \sum_{i=1}^{\infty} \frac{J^i y}{\lambda^i} \qquad \forall y \in L^2(\Omega).$$

DIM. Per la (3) di pag.1 l'ipotesi su λ comporta che $|\lambda| > \|J\|$. Per la (3.3) di pag.216 ne segue che esiste l'operatore inverso $(J-\lambda I)^{-1}$, è definito in tutto lo spazio $L^2(\Omega)$, ed è dato dalla relazione:

$$(J-\lambda I)^{-1} = -\frac{1}{\lambda} I - \frac{1}{\lambda} \sum_{i=1}^{\infty} \frac{J^i}{\lambda^i},$$

che per la prima delle (13) si scrive:

$$(20) \qquad (J-\lambda I)^{-1} = -\frac{1}{\lambda} I - \frac{1}{\lambda} J^{(\lambda)}.$$

In virtù della (14), dalla (20) si deduce la (19).

Con riferimento all'equazione integrale (4), il teorema dianzi acquisito si può enunciare come segue, tenuto conto delle (14), (15):

(3.2). Se $|\lambda| > \|K\|$, l'equazione integrale $Jx - \lambda x = y$, ossia:

$$(21) \qquad \int_\Omega K(\sigma,t)x(t)dt - \lambda x(\sigma) = y(\sigma)$$

ammette, qualunque sia il termine noto $y \in L^2(\Omega)$, una e una sola soluzione $x_\lambda \in L^2(\Omega)$, data dalla formula:

$$(22) \qquad x_\lambda = -\frac{y}{\lambda} - \frac{1}{\lambda}\sum_{\tau=1}^{\infty} \frac{J^\tau y}{\lambda^\tau} .$$

La funzione x_λ ha l'espressione:

$$(23) \qquad x_\lambda(\sigma) = -\frac{y(\sigma)}{\lambda} - \frac{1}{\lambda}\int_\Omega K^{(\lambda)}(\sigma,t)y(t)dt \quad \text{q.o. in } \Omega .$$

La funzione (23) si può esprimere, come andiamo a dimostrare, mediante la serie di Neumann relativa al nucleo K [pag.4 e (7) pag.2], cioè nella forma:

$$(24) \qquad x_\lambda(\sigma) = -\frac{y(\sigma)}{\lambda} - \frac{1}{\lambda}\int_\Omega \sum_{\tau=1}^{\infty} \frac{K_{\tau-1}(\sigma,t)}{\lambda^\tau} y(t)dt \quad \text{q.o. in } \Omega .$$

Si tratta di stabilire l'uguaglianza:

$$(25) \qquad K^{(\lambda)}(\sigma,t) = \sum_{\tau=1}^{\infty} \frac{K_{\tau-1}(\sigma,t)}{\lambda^\tau} \qquad \text{q.o. in } \Omega\times\Omega$$

che non è diretta conseguenza della seconda delle (13), dato che quest'ultima va intesa nel senso di $L^2(\Omega\times\Omega)$.

DIM. Identificando $\forall \tau \in N$ l'elemento $K_{\tau-1}$ di $L^2(\Omega\times\Omega)$ con un suo rappresentante definito in tutto $\Omega\times\Omega$, proveremo che $\forall (\sigma,t) \in \Omega\times\Omega$ la serie numerica a secondo membro della (25) è convergente. Dopo ciò, per definizione di somma di una serie di funzioni, il secondo membro della (25) sarà il valore nel punto (σ,t) della funzione $(\sum_{\tau=1}^{\infty} K_{\tau-1}/\lambda^\tau)$: essendo quest'ultima l'elemento di $L^2(\Omega\times\Omega)$ che abbiamo denotato con $K^{(\lambda)}$, ne seguirà la (25).

Per ogni $(\sigma,t) \in \Omega\times\Omega$ poniamo:

$$\|K(\sigma,\cdot)\|_{L^2(\Omega)} = f(\sigma) \quad , \quad \|K(\cdot,t)\|_{L^2(\Omega)} = g(t)$$

e dimostriamo preliminarmente che:

$$(26) \qquad \|K_{\tau-1}(\cdot,t)\|_{L^2(\Omega)} \le g(t)\|K\|^{\tau-1} \qquad \forall \tau \in N .$$

Siccome $K_0 = K$, la (26) è banale per $\tau = 1$. Procedendo per induzione, proviamo che se è vera la (26), allora è vera anche la relazione che si ottiene sostituendo in essa $\tau-1$ con τ. Ricordando la (7) di pag.2, per la disuguaglianza di Schwartz e per la (26) si ha:

$$|K_\tau(\sigma,t)|^2 = |(K(\sigma,\cdot),K_{\tau-1}(\cdot,t))|^2 \le f^2(\sigma)\|K_{\tau-1}(\cdot,t)\|^2 \le f^2(\sigma)g^2(t)\|K\|^{2\tau-2} .$$

Integrando rispetto ad σ, siccome l'integrale di $f^2(\sigma)$ è uguale a $\|K\|^2$, se ne deduce la (26) con τ in luogo di $\tau-1$.

Tenuto conto della (26) [con $\tau-1$ in luogo di τ], si ha per $\tau \ge 2$:

$$\frac{|K_{\tau-1}(\sigma,t)|}{|\lambda|^\tau} = \frac{1}{|\lambda|^\tau}|(K(\sigma,\cdot),K_{\tau-2}(\cdot,t))| \le \frac{1}{|\lambda|^\tau}\|K(\sigma,\cdot)\|\cdot\|K_{\tau-2}(\cdot,t)\| \le \frac{f(\sigma)g(t)}{|\lambda|^2}\left(\frac{\|K\|}{|\lambda|}\right)^{\tau-2} .$$

Ne segue che la serie a secondo membro della (25) è convergente, dato che la serie dei moduli è maggiorata (dal secondo termine in poi) da una serie convergente (essendo $\|K\|/|\lambda| < 1$).

Vogliamo ora indicare un'altra dimostrazione della (3.2), che consente di "costruire" la soluzione x_λ della (21) con un procedimento iterativo.

Se l'equazione integrale $Jx-\lambda x=y$ si scrive nella forma $Jx/\lambda - y/\lambda = x$, e si denota con $F(x)$ il primo membro di questa, l'equazione diviene:

$$F(x)=x\ .$$

Perciò il problema equivale a ricercare un punto fisso della funzione:

$$F : z\in L^2(\Omega) \to (-y/\lambda + Jz/\lambda) \in L^2(\Omega)\ .$$

Orbene, in virtù della (2) di pag.2, qualunque siano $z_1, z_2 \in L^2(\Omega)$ si ha:

$$\|F(z_1)-F(z_2)\| = \frac{1}{|\lambda|}\|J(z_1-z_2)\| \le \frac{\|K\|}{|\lambda|}\|z_1-z_2\|\ ,$$

e quindi, essendo $|\lambda|>\|K\|$, la F è una contrazione dello spazio $L^2(\Omega)$.

Per la (1.2) di pag.285 ne segue che F ammette uno ed un solo punto fisso x_λ, che è il limite della successione $(x_n)_{n\in N}$ ottenuta scegliendo ad arbitrio un punto $x_0\in L^2(\Omega)$ e ponendo $x_n = F(x_{n-1})\ \forall n\in N$.

Scelto $x_0=-y/\lambda$, si verifica subito che risulta:

$$x_n = -\frac{y}{\lambda} - \frac{1}{\lambda}\sum_{i=1}^{n}\frac{J^i y}{\lambda^i} \tag{27}$$

e passando al limite si ritrova la (22).

DIM. (27). La (27) è vera per $n=1$: infatti, essendo $x_0=-y/\lambda$, il secondo membro per $n=1$ è uguale ad $F(x_0)$. Procedendo per induzione, ammettiamo che il secondo membro della (27) sia uguale ad $F(x_{n-1})$, e proviamo che $F(x_n)$ è uguale al secondo membro della (27) con $n+1$ in luogo di n. Ciò consegue dalle uguaglianze:

$$F(x_n)+y/\lambda = Jx_n/\lambda = -Jy/\lambda^2 - \frac{1}{\lambda^2}\sum_{i=1}^{n}\frac{J^{i+1}y}{\lambda^i} = -\frac{1}{\lambda}\sum_{i=0}^{n}\frac{J^{i+1}y}{\lambda^{i+1}} = -\frac{1}{\lambda}\sum_{i=1}^{n+1}\frac{J^i y}{\lambda^i}\ .$$

Per la citata (1.2) di pag.285 risulta inoltre:

$$\|x_n - x_\lambda\|_{L^2(\Omega)} \le \frac{\|K\|\cdot\|y\|}{|\lambda|(|\lambda|-\|K\|)}\left(\frac{\|K\|}{|\lambda|}\right)^n \qquad \forall n\in N. \tag{28}$$

La (28) fornisce una valutazione dell'errore, in media quadratica, che si commette assumendo x_n come soluzione dell'equazione.

Si dimostra, ma noi vi rinunciamo, che se esiste una costante c tale che $\|K(s,\cdot)\|_{L^2(\Omega)}\le c$ q.o. in Ω, risulta $\lim x_n = x_\lambda$ uniformemente in Ω.

4. In questo numero supponiamo che il nucleo $K(s,t)$ sia simmetrico, sicchè l'operatore compatto J è anche autoaggiunto.

Ci riferiremo al caso che lo spettro sia numerabile; peraltro le considerazioni che seguono sono valide anche quando lo spettro è finito,

con ovvie varianti formali tra cui quella che alle serie subentrano somme con un numero finito di addendi.

Lo spettro di J è costituito dal punto 0 e dagli autovalori, che sono tutti reali; il punto 0, se non è autovalore, appartiene allo spettro continuo [(5.4) pag. 245].

Il teorema di Hilbert-Schmidt [pag. 246] fornisce la rappresentazione spettrale dell'operatore integrale J. Precisamente, introdotta la successione $\{\lambda_n\}$ degli autovalori e una successione ortonormale $\{x_n\}$ di autovettori corrispondenti [elementi di $L^2(\Omega)$] come a pag. 245 [inizio n. 4, con $A = J$], risulta:

$$Jx = \sum_{n=1}^{\infty} \lambda_n x_n \int_\Omega x(t)\bar{x}_n(t)\,dt \qquad \forall x \in L^2(\Omega). \tag{29}$$

Se il punto 0 non è autovalore di J si ha poi [(20) pag. 247]:

$$x = \sum_{n=1}^{\infty} x_n \int_\Omega x(t)\bar{x}_n(t)\,dt \qquad \forall x \in L^2(\Omega). \tag{30}$$

Il secondo membro della (30) è la serie di Fourier del punto x rispetto al sistema ortonormale $\{x_n\}$ [pag. 127 e (6.3) pag. 126], e perciò nell'ipotesi che 0 non sia autovalore di J il sistema ortonormale $\{x_n\}$ è completo [(6.5) pag. 129]. In tale ipotesi è facile verificare che risulta:

$$\sum_{n=1}^{\infty} \lambda_n^2 = \|K\|^2 \quad ; \tag{31}$$

peraltro la (31) sussiste anche se 0 è autovalore di J.

DIM. Per semplicità ci limitiamo al caso che 0 non sia autovalore di J, sicché per $K(s,\cdot)$, elemento di $L^2(\Omega)$ [teor. di Fubini, nota (1) pag. 169], sussiste l'uguaglianza di Parseval [(6.5) pag. 129 e (13) pag. 128]. Risulta quindi:

$$\sum_{n=1}^{\infty} |(K(s,\cdot), x_n)|^2 = \|K(s,\cdot)\|^2. \tag{32}$$

D'altra parte, essendo x_n autovettore di J relativo all'autovalore λ_n, si ha $Jx_n = \lambda_n x_n$, ossia:

$$\lambda_n x_n(s) = \int_\Omega K(s,t)x_n(t)\,dt = (K(s,\cdot), \bar{x}_n).$$

La (32) diviene allora

$$\sum_{n=1}^{\infty} |\lambda_n x_n(s)|^2 = \int_\Omega |K(s,t)|^2\,dt.$$

Questa uguaglianza sussiste per s quasi ovunque in Ω. Integrando si ottiene:

$$\int_\Omega \sum_{n=1}^{\infty} |\lambda_n x_n(s)|^2\,ds = \|K\|^2.$$

Per il teorema di Beppo Levi [nota (1) pag. 335] nel primo membro si può trasportare il simbolo sommatorio fuori dell'integrale. Si ha allora:

$$\sum_{n=1}^{\infty} |\lambda_n|^2 \|x_n\|^2 = \|K\|^2$$

e ciò prova la (31), dato che $\|x_n\| = 1 \ \forall n \in N$.

5. In questo numero riprendiamo l'equazione integrale

(33) $$(J-\lambda I)x = y \ ,$$

ossia l'equazione:

(33') $$\int_\Omega K(s,t)x(t)\,dt - \lambda x(s) = y(s) \quad \text{q.o. in } \Omega$$

introdotta a pag. 2. Nell'ipotesi $|\lambda| > \|K\|$ abbiamo stabilito il teorema di esistenza e unicità (3.2) [pag. 7], che fornisce due espressioni della soluzione; da notare che nella prima [cfr. (22)] figura la somma di una serie che va intesa nel senso di $L^2(\Omega)$, sicché l'uguaglianza esprime la soluzione x_λ, ma non $x_\lambda(s)$ per s quasi ovunque in Ω come accade per la (23). Abbiamo poi constatato che la (23) si può scrivere nella forma (24), mediante la serie di Neumann relativa al nucleo K.

Siamo ritornati sull'equazione (33), o (33'), perché qualunque sia $y \in L^2(\Omega)$ essa ammette una e una sola soluzione non soltanto per $|\lambda| > \|K\|$, ma per ogni $\lambda \neq 0$ che non sia autovalore di J: invero ogni λ siffatto, per quanto si è detto a pag. 6, appartiene all'insieme risolvente di J.

Vogliamo occuparci brevemente della (33) anche per gli altri punti λ dell'insieme risolvente di J. Per completare l'argomento prenderemo poi in considerazione anche il caso che nella (33) il numero λ sia autovalore di J (diverso da 0).

La citata (23) esprime la soluzione dell'equazione, per $|\lambda| > \|K\|$, mediante l'operatore integrale $J^{(\lambda)}$ generato dal nucleo $K^{(\lambda)}(s,t)$ [cfr. (15) pag. 4]. Orbene, anche negli altri casi la soluzione si esprime nella forma (23): all'operatore integrale generato dal nucleo $K^{(\lambda)}(s,t)$ subentra l'operatore integrale generato da un altro nucleo $R^{(\lambda)}(s,t)$, il quale ha significato per qualunque $\lambda \neq 0$ che non sia autovalore di J, e coincide con $K^{(\lambda)}(s,t)$ per $|\lambda| > \|K\|$.

Abbiamo così illustrato in via preliminare il teorema seguente, del quale non riportiamo la dimostrazione.

(5.1). Per ogni $\lambda \in \mathbb{C}-\{0\}$ che non sia autovalore di J, esiste una funzione $R^\lambda(\sigma,t)$ elemento di $L^2(\Omega\times\Omega)$, che consente di esprimere la soluzione x_λ dell'equazione integrale (33), (33') nella forma:

$$(35) \qquad x_\lambda(\sigma) = -\frac{y(\sigma)}{\lambda} - \frac{1}{\lambda}\int_\Omega R^{(\lambda)}(\sigma,t)\,y(t)\,dt \qquad \text{q.o. in } \Omega .$$

La funzione $R^\lambda(\sigma,t)$ prende il nome di nucleo risolvente dell'operatore J. Aggiungiamo che se il nucleo $K(\sigma,t)$ è simmetrico (sicchè J è autoaggiunto), per ogni $\lambda\neq 0$ che non sia autovalore di J la soluzione x_λ della (33) si può esprimere, utilizzando una conseguenza del teorema di Hilbert-Schmidt, mediante gli autovalori di J e certi autovettori ad essi corrispondenti. Precisamente, con le stesse notazioni utilizzate per la rappresentazione spettrale di J [pag. 9], dal teor. (4.3) di pag. 248 si deduce:

$$(36) \qquad x_\lambda = \frac{-y}{\lambda} - \frac{1}{\lambda}\sum_{n=1}^{\infty} \frac{\lambda_n}{\lambda-\lambda_n}\, x_n \int_\Omega y(t)\,\overline{x}_n(t)\,dt \;,$$

dove la serie, naturalmente, va intesa nel senso di $L^2(\Omega)$.

Ancora nell'ipotesi $K(\sigma,t)$ simmetrico, si dimostra (ma noi vi rinunciamo) che per $\lambda\neq 0$ e $\lambda\neq\lambda_n$ $\forall n\in N$ il nucleo risolvente $R^{(\lambda)}(\sigma,t)$ ha l'espressione:

$$(37) \qquad R^{(\lambda)}(\sigma,t) = \sum_{n=1}^{\infty} \frac{\lambda_n}{\lambda-\lambda_n}\, x_n(\sigma)\,\overline{x}_n(t) \qquad \text{q.o. in } \Omega\times\Omega .$$

Notiamo che se, procedendo formalmente, si scrive l'espressione di $x_\lambda(\sigma)$ q.o. in Ω ottenuta dalla (36) con $y(\sigma)$ in luogo di y ed $x_n(\sigma)$ in luogo di x_n, scambiando poi le operazioni di integrazione e di "sommazione" si ottiene:

$$x_\lambda(\sigma) = -\frac{y(\sigma)}{\lambda} - \frac{1}{\lambda}\int_\Omega \left[\sum_{n=1}^{\infty} \frac{\lambda_n}{\lambda-\lambda_n}\, x_n(\sigma)\,\overline{x}_n(t)\right] y(t)\,dt \qquad \text{q.o. in } \Omega .$$

Orbene, tale relazione effettivamente sussiste: basta sostituire l'espressione (37) nella (35).

* * *

Prima di concludere ci sembra opportuno che non resti sottinteso, per l'equazione integrale (33), il caso che λ sia un autovalore non nullo di J: allora l'equazione per certi $y\in L^2(\Omega)$ non ammette alcuna soluzione, e per certi $y\in L^2(\Omega)$ ammette infinite soluzioni.

Consideriamo dunque l'equazione integrale (con nucleo non necessariamente simmetrico):

(38) $\int_\Omega K(s,t)x(t)dt - \lambda_0 x(s) = y(s)$ q.o. in Ω

dove λ_0 e' un autovalore non nullo dell'operatore integrale J generato dal nucleo $K(s,t)$. Siccome l'equazione omogenea associata alla (38) [ossia la (38) con $y \equiv 0$] ammette infinite soluzioni (che sono gli autovettori corrispondenti all'autovalore λ_0), delle due eventualita' previste nel teorema dell'alternativa [(6.2) pag. 235] si verifica la seconda: se ne deduce quanto segue.

Supposto che l'autovalore λ_0 abbia molteplicita' $h \geq 1$, e scelte h soluzioni linearmente indipendenti dell'equazione

(39) $\int_\Omega \overline{K(t,s)}\, u(t)\, dt - \bar{\lambda}_0 u(s) = 0$,

diciamole $u_1, \dots, u_h$, la (38) ammette soluzioni se e solo se il termine noto y soddisfa alle condizioni:

(40) $\int_\Omega y(t) u_1(t) dt = 0, \dots, \int_\Omega y(t) u_h(t) dt = 0$.

Considerando l'equazione aggiunta della (38), la (39) e' l'equazione omogenea associata ad essa. Le funzioni $u_1, \dots, u_h$ costituiscono una base del nucleo dell'operatore $J^* - \bar{\lambda}_0 I$, aggiunto di $J - \lambda_0 I$; le (40) esprimono che il termine noto y dell'equazione (38) e' ortogonale al nucleo predetto.

Quando le (40) sono soddisfatte, l'equazione (38) ammette infinite soluzioni. Detti $x_1, \dots, x_h$ degli autovettori linearmente indipendenti relativi all'autovalore λ_0 (in numero di h, molteplicita' dell'autovalore), le soluzioni della (38) sono date dalla formula:

$$x = c_1 x_1 + \dots + c_h x_h + x_0,$$

dove x_0 e' una soluzione della stessa (38) e $c_1, \dots, c_h$ sono costanti arbitrarie.

* * *

Segnaliamo al lettore l'esercizio svolto a pag. 237, al quale si aggiungono le considerazioni svolte a pag. 251.

EQUAZIONI INTEGRALI DI FREDHOLM

CON NUCLEO DI CLASSE $N^{(\alpha)}$

Vogliamo dare un brevissimo cenno di certi operatori integrali che si incontrano frequentemente nella Teoria del potenziale, e delle equazioni integrali ad essi relative, il cui nucleo non appartiene necessariamente ad $L^2(\Omega\times\Omega)$.

Sia Ω un aperto limitato di R^m, il cui diametro indicheremo con δ, e sia $K(s,t)$ una funzione definita in ogni punto $(s,t)\in\bar\Omega\times\bar\Omega$ con $s\neq t$. Fissato un numero positivo α, si dice che K è un <u>nucleo di classe $N^{(\alpha)}$</u> quando K è continua in ogni punto del suo insieme di definizione, e soddisfa alla seguente condizione:

— se $0<\alpha<m$, esiste $c\in R$ tale che:

$$|K(s,t)|\leq \frac{c}{|s-t|^{m-\alpha}} \qquad \forall (s,t)\in\bar\Omega\times\bar\Omega,\ s\neq t$$

— se $\alpha=m$, esiste $c\in R$ tale che:

$$|K(s,t)|\leq c\log\frac{2\delta}{|s-t|} \qquad \forall (s,t)\in\bar\Omega\times\bar\Omega,\ s\neq t$$

— se $\alpha>m$, K è prolungabile per continuità su tutto $\bar\Omega\times\bar\Omega$.

Supposto K di classe $N^{(\alpha)}$ e fissato $x\in L^1(\Omega)$, si dimostra che per ogni $s\in\Omega$ la funzione $K(s,\cdot)x$ è sommabile in Ω, e la funzione y definita ponendo:

$$y(s)=\int_\Omega K(s,t)x(t)\,dt \qquad \forall s\in\Omega \tag{41}$$

appartiene ad $L^1(\Omega)$.

<u>L'operatore lineare che ad $x\in L^1(\Omega)$ associa la funzione $y\in L^1(\Omega)$ definita dalla (41) è continuo</u>, e si chiama <u>l'operatore integrale generato dal nucleo K</u>:

lo indicheremo con J, anche se il nucleo K non appartiene necessariamente ad $L^2(\Omega\times\Omega)$.

Siccome lo spazio duale di $L^1(\Omega)$ è equivalente ad $L^\infty(\Omega)$ [es. 4 pag. 181 con $p=1$ e quindi $p'=+\infty$], identificando $(L^1(\Omega))^*$ con $L^\infty(\Omega)$ si ha che l'aggiunto J^* di J [n. 5 pag. 227.5] è un operatore di $L^\infty(\Omega)$ in $L^\infty(\Omega)$: precisamente, ad $u\in L^\infty(\Omega)$ l'operatore J^* associa la funzione v definita ponendo:

$$v(s)=\int_\Omega \overline{K(t,s)}\,u(t)\,dt \qquad \forall s\in\Omega$$

che, come si può dimostrare, appartiene ad $L^\infty(\Omega)$.

Se si considera, per ogni $y\in L^1(\Omega)$, l'equazione integrale nell'incognita $x\in L^1(\Omega)$:

(42) $$Jx-\lambda x=y \qquad \lambda\in C-\{0\},$$

l'aggiunta della (42) è allora, per ogni $u'\in L^\infty(\Omega)$, l'equazione nella incognita $u\in L^\infty(\Omega)$:

(43) $$J^*u-\bar\lambda u=u'.$$

Orbene, pur essendo J non necessariamente compatto [dato che il nucleo K non appartiene necessariamente ad $L^2(\Omega\times\Omega)$], si dimostra che:

<u>Se il nucleo K è di classe $N^{(a)}$, per le equazioni (42), (43) sussiste la tesi del teorema dell'alternativa</u> [pag. 233].

Pertanto, o le equazioni omogenee associate alle (42), (43) ammettono entrambe soltanto la soluzione nulla, ed allora le (42), (43) ammettono, qualunque siano i termini noti, una e una sola soluzione; oppure il numero λ è autovalore di J, $\bar\lambda$ è autovalore di J^*, e $\lambda, \bar\lambda$ hanno la stessa molteplicità.

Nella seconda delle predette eventualità, per l'equazione (42)

si può ripetere quanto detto con $\lambda=\lambda_0$ nella precedente pagina 12 a proposito dell'equazione (38) [nella quale il nucleo K appartiene ad $L^2(\Omega\times\Omega)$], con le necessarie varianti che andiamo a precisare.

La (38) è "ambientata" nello spazio $L^2(\Omega)$ [spazio di Hilbert, quindi isomorfo al suo duale], nel senso che a tale spazio appartiene il termine noto e nello spazio stesso si cercano le soluzioni; anche la (39), omogenea associata all'equazione aggiunta della (38), naturalmente è ambientata in $L^2(\Omega)$; utilizzando la notazione $(\cdot,\cdot)$ per il prodotto scalare in $L^2(\Omega)$, le condizioni (40) si scrivono:

$$(44)\qquad (u_1,y)=0\ ,\ (u_2,y)=0\ ,\ \dots\ ,\ (u_m,y)=0\ .$$

Orbene, l'equazione (42) è la stessa (38), ambientata però nello spazio $L^1(\Omega)$, e l'omogenea associata alla (43) è la (39), ambientata però nello spazio duale di $L^1(\Omega)$, identificato con $L^\infty(\Omega)$. Dunque se il nucleo $K(s,t)$ è di classe $N^{(1)}$, sussiste quanto scritto a pag. 12 salvo a cambiare gli spazi come appena detto; naturalmente nel caso in questione le condizioni (40) non si esprimono mediante le (44): gli elementi $u_1,\dots,u_m$ di $L^\infty(\Omega)$ si identificano con dei funzionali definiti in $L^1(\Omega)$ [elementi del duale di $L^1(\Omega)$], e alle (44) subentrano le relazioni analoghe che si ottengono sostituendo la notazione $(\cdot,\cdot)$ del prodotto scalare in $L^2(\Omega)$ con la dualità $\langle\cdot,\cdot\rangle$ tra $L^1(\Omega)$ ed $(L^1(\Omega))^*$ [pag. 171].

Nell'ipotesi posta sul nucleo K l'operatore integrale J gode di importanti proprietà. Allo scopo di formularne alcune è opportuno ricordare che l'aperto Ω è un sottoinsieme di R^m, sicché gli elementi di $L^p(\Omega)$ [$p\in[1,+\infty]$] sono funzioni di m variabili, e il nucleo

$K(s,t)$ di classe $N^{(\nu)}$ è funzione di $2m$ variabili.

Orbene, come nella definizione della classe $N^{(\nu)}$, anche nelle proposizioni che seguono intervengono i numeri α ed m, dato che bisogna distinguere i tre casi $\alpha<m$, $\alpha=m$, $\alpha>m$.

1. Se $\alpha<m$, per ogni $p\in[1, m/\alpha[$ si ha:

$$x\in L^p(\Omega) \Longrightarrow Jx\in L^{r(p)}(\Omega) \quad \text{con } r(p)=mp/(m-\alpha p)$$

e per gli altri valori di p (ossia $\geq m/\alpha$) si ha:

(45) $$x\in L^{m/\alpha}(\Omega)\Longrightarrow Jx\in L^r \ \ \forall r\in[1,+\infty[\ ; \ x\in L^p(\Omega) \text{ con } p>\frac{m}{\alpha} \Longrightarrow Jx\in C^0(\Omega).$$

2. Se $\alpha=m$ vale ancora la prima delle (45), cioè si ha:

$$x\in L^1(\Omega)\Longrightarrow Jx\in L^r(\Omega) \ \ \forall r\in[1,+\infty[.$$

3. Se $\alpha>m$ si ha:

(46) $$x\in L^1(\Omega)\Longrightarrow Jx\in C^0(\Omega).$$

Con riferimento a ciascuna delle precedenti implicazioni, l'operatore lineare $x\longrightarrow Jx$, definito nello spazio a sinistra dell'implicazione ed a valori nello spazio a destra, è continuo; in altri termini [(2.1) pag. 138] esiste una costante c tale che $\|Jx\|\leq c\|x\|$ per ogni punto x dello spazio di definizione [naturalmente la norma di x va intesa in quest'ultimo spazio, mentre quella di Jx va intesa nello spazio di arrivo]. Con riferimento alla (46) l'operatore $x\to Jx$ è anche compatto.

Si può dire che ciascuno dei predetti operatori lineari trasforma ogni punto x dello spazio di definizione in un punto di uno spazio "qualitativamente migliore" dato che, essendo Ω limitato, la famiglia di spazi $(L^p(\Omega))_{p\in[1,+\infty]}$ è decrescente rispetto all'inclusione. Notiamo che tutti gli spazi in questione sono inclusi in $L^1(\Omega)$.

FORME DIFFERENZIALI DI GRADO K (1)

Il duale di R^n sarà denotato con $\overset{+}{R}{}^n$. Un elemento di tale spazio (forma lineare in R^n):

$$f : (t_1, \dots, t_n) \in R^n \longrightarrow c_1 t_1 + \cdots + c_n t_n$$

può essere identificato con la n-pla $(c_1, \dots, c_n)$ dei suoi coefficienti. Per i nostri scopi è opportuno identificare il numero $f(t_1, \dots, t_n)$ con la matrice (ad una riga ed una colonna) di cui è l'unico elemento. Se allora si considera la matrice:

$$a = [c_1, \dots, c_n]$$

ad una riga ed n colonne (vettore riga), nonché la matrice:

$$v = \begin{bmatrix} t_1 \\ \vdots \\ t_n \end{bmatrix}$$

ad n righe ed una colonna (vettore colonna), si può scrivere $f(v) = av$.

Perciò noi identificheremo lo spazio R^n con lo spazio dei vettori colonna, e il suo duale $\overset{+}{R}{}^n$ con lo spazio dei vettori riga. Di conseguenza una k-pla $(v_1, \dots, v_k)$ di elementi di R^n, ossia di vettori colonna, sarà identificata con la matrice:

(1) $$[v_1, \dots, v_k]$$

ad n righe, e una k-pla $(a_1, \dots, a_k)$ di elementi di $\overset{+}{R}{}^n$, ossia di vettori riga, sarà identificata con la matrice:

(2) $$\begin{bmatrix} a_1 \\ \vdots \\ a_k \end{bmatrix}$$

ad n colonne. Abbiamo indicato con n e k degli elementi di N arbitrari.

Dopo ciò una funzione di k variabili, se definita in $(R^n)^k$ opera su matrici del tipo (1), se definita in $(\overset{+}{R}{}^n)^k$ opera su matrici del tipo (2).

Detto S uno spazio vettoriale reale, consideriamo ora una funzione reale F di k variabili, definita in S^k $(k \geq 2)$. Se F è lineare rispetto a ciascuna delle variabili, si dice che F è una <u>forma multilineare in S</u>, <u>di grado k</u>, oppure una <u>forma k-lineare in S</u> (<u>bilineare</u> per $k=2$). Una tale funzione F si dice poi <u>alternante</u> se il suo valore in un qualunque punto cambia di segno quando si scambiano tra loro due coordinate del punto.

A noi occorre prendere in considerazione forme k-lineari alternanti nello spazio R^n e nel suo duale $\overset{+}{R}{}^n$.

* * *

(1) Per le dimostrazioni si rimanda al testo ANALISI MATEMATICA DUE di Fusco-Marcellini-Sbordone, dal quale questi appunti sono stati tratti. Si veda anche FUNCTIONS OF SEVERAL VARIABLES di Fleming.

Una forma K-lineare alternante in R^n [in $\overset{+}{R}{}^n$] ($K \geq 2$):

$$\omega : (R^n)^K \longrightarrow R \qquad [\overset{+}{\omega} : (\overset{+}{R}{}^n)^K \longrightarrow R]$$

viene chiamata <u>covettore di grado K</u>, o <u>K-covettore</u> [<u>vettore di grado K</u>, o <u>K-vettore</u>] <u>di R^n</u>; il riferimento ad R^n sarà sottinteso, dato che supponiamo di aver fissato $n \geq 2$ una volta per tutte. Quando non occorre specificare il grado, useremo semplicemente la parola <u>covettore</u> [<u>vettore</u>].

Dunque <u>i K-covettori operano su matrici a K-colonne</u>, che scriveremo nella forma (1) evidenziandone le colonne, mentre <u>i K-vettori operano su matrici a K righe</u>, che scriveremo nella forma (2) evidenziandone le righe. Per uniformità di linguaggio si attribuisce a K anche il valore 1, intendendo per <u>1-covettore</u> [per <u>1-vettore</u>] una forma lineare che opera sui vettori colonna [sui vettori riga].

L'insieme costituito dai K-covettori [dai K-vettori] sarà denotato con $\Omega_K(R^n)$ [con $\overset{+}{\Omega}_K(R^n)$], o semplicemente con Ω_K [con $\overset{+}{\Omega}_K$], e si supporrà munito della struttura di spazio vettoriale reale con le usuali operazioni di addizione e di moltiplicazione per uno scalare.

Introduciamo ora alcuni covettori che ci occorreranno nel seguito.

Per ogni K-pla $\tau = (\tau_1, \dots, \tau_K)$ di interi appartenenti all'insieme $\{1, 2, \dots, n\}$, interi che chiameremo <u>indici</u>, con il simbolo:

$$(3) \qquad dx_{\tau_1} \wedge dx_{\tau_2} \wedge \dots \wedge dx_{\tau_K}$$

si denota il K-covettore che ad ogni matrice:

$$A = [v_1, \dots, v_K] = \begin{bmatrix} t_{11} & t_{12} & \dots & t_{1K} \\ t_{21} & t_{22} & \dots & t_{2K} \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ t_{n1} & t_{n2} & \dots & t_{nK} \end{bmatrix}$$

associa il numero:

$$\det \begin{bmatrix} t_{\tau_1 1} & t_{\tau_1 2} & \dots & t_{\tau_1 K} \\ t_{\tau_2 1} & t_{\tau_2 2} & \dots & t_{\tau_2 K} \\ & & & \\ t_{\tau_K 1} & t_{\tau_K 2} & \dots & t_{\tau_K K} \end{bmatrix}$$

Il simbolo (3) si abbrevia scrivendo:

$$dx_{\tau_1, \tau_2, \dots, \tau_K} \quad \text{oppure} \quad dx_\tau \ .$$

Qualunque sia l'intero m tale che $1<m<\kappa$, posto $\tau'=(\tau_1,\dots,\tau_{m-1})$, $\tau''=(\tau_m,\dots,\tau_\kappa)$ il simbolo dx_τ si scrive anche

$$dx_{\tau'}\wedge dx_{\tau''} \quad \text{oppure} \quad dx_{\tau',\tau''} .$$

Nel caso che τ sia un indice, ossia un intero $i\in\{1,\dots,n\}$, il simbolo $dx_\tau = dx_i$ rappresenta un particolare 1-covettore: precisamente la corrispondenza che ad ogni vettore colonna $[v]$ associa la componente i-ma del vettore v di R^n. Naturalmente il covettore $dx_{1,2,\dots,n}$, di grado n, associa ad ogni matrice $A=[v_1,\dots,v_n]$, ad n righe, il determinante di A.

Ovviamente se nella κ-pla $\tau=(\tau_1,\dots,\tau_\kappa)$ vi sono due indici uguali (il che deve necessariamente accadere nel caso $\kappa>n$), il covettore dx_τ è la funzione identicamente nulla in $(R^n)^\kappa$.

Denotiamo con I_κ l'insieme costituito dalle κ-ple $\tau=(\tau_1,\dots,\tau_\kappa)$ con $\kappa\le n$, ed inoltre se $\kappa>1$ con $\tau_1<\tau_2<\dots<\tau_\kappa$: al variare di τ in I_κ il κ-covettore dx_τ descrive un sottoinsieme di Ω_κ che costituisce, come facilmente si verifica, una base dello spazio vettoriale Ω_κ (<u>base canonica</u>). Di conseguenza Ω_κ ha dimensione $\binom{n}{\kappa}$ per $\kappa\le n$, mentre per $\kappa>n$ si riduce al solo κ-covettore nullo.

Dunque ogni κ-covettore ω, per $\kappa\le n$, è una combinazione lineare dei κ-covettori dx_τ con $\tau\in I_\kappa$, i coefficienti essendo univocamente determinati. Denotando con ω_τ il coefficiente di dx_τ, che scriveremo anche nella forma $\omega_{\tau_1,\dots,\tau_m}$, si ha allora:

$$\omega=\sum_{\tau\in I_\kappa}\omega_\tau\,dx_\tau .$$

Pertanto il corrispondente mediante ω di una matrice $A=[v_1,\dots,v_\kappa]$ è una combinazione lineare dei determinanti minori di ordine massimo (ossia di ordine κ) estratti da A (sicché è 0 se il rango di A è minore di κ).

Ovviamente il covettore ω si può anche esprimere come somma di tutti i covettori dx_τ (cioè non soltanto di quelli della base canonica, con $\tau\in I_\kappa$), scrivendo:

$$\omega=\sum_{\sigma_1,\dots,\sigma_\kappa}^{1,\dots,n}\omega_{\sigma_1,\dots,\sigma_\kappa}\,dx_{\sigma_1,\dots,\sigma_\kappa} ,$$

ed allora per ogni $\tau\in I_\kappa$ il coefficiente ω_τ di dx_τ è la somma di $\kappa!$ addendi.

Nello spazio Ω_κ si introduce il prodotto scalare ponendo, qualunque siano i covettori $\omega=\sum_{\tau\in I_\kappa}\omega_\tau dx_\tau$, $\eta=\sum_{\tau\in I_\kappa}\eta_\tau dx_\tau$:

$$(\omega,\eta)=\sum_{\tau\in I_\kappa}\omega_\tau\eta_\tau$$

e quindi la norma di ω in Ω_κ è data da:

$$\|\omega\|=(\omega,\omega)^{1/2}=\Big[\sum_{\tau\in I_\kappa}\omega_\tau^2\Big]^{1/2} .$$

Nell'insieme costituito da tutti i κ-covettori con $\kappa\in N$ si introduce l'operazione di <u>molti-</u>

plicazione esterna. Si chiama prodotto esterno dei covettori:

$$\omega = \sum_{\tau \in I_{n,k}} \omega_\tau \, dx_\tau \quad , \quad \omega' = \sum_{\sigma \in I_{n,h}} \omega'_\sigma \, dx_\sigma \quad ,$$

rispettivamente di gradi k ed h, e si denota con $\omega \wedge \omega'$, il covettore:

$$\omega \wedge \omega' = \sum_{\substack{\tau \in I_{n,k} \\ \sigma \in I_{n,h}}} \omega_\tau \omega'_\sigma \, dx_\tau \wedge dx_\sigma = \sum_{\substack{\tau \in I_{n,k} \\ \sigma \in I_{n,h}}} \omega_\tau \omega'_\sigma \, dx_{\tau,\sigma}$$

di grado $k+h$. Si noti che la $(k+h)$-pla $p = (\tau,\sigma) = (\tau_1,\dots,\tau_k,\sigma_1,\dots,\sigma_h)$ non è necessariamente un elemento dell'insieme $I_{n,k+h}$, perchè non è detto che le coordinate siano disposte in ordine crescente; si noti altresì che nel caso $k+h > n$ il prodotto esterno è nullo [ossia è la funzione identicamente nulla in $(R^n)^{k+h}$].

Ovviamente l'operazione di moltiplicazione esterna è associativa, e non è commutativa in quanto risulta:

$$\omega \wedge \omega' = (-1)^{kh} \, \omega' \wedge \omega \, .$$

Inoltre la predetta operazione (funzione delle due variabili ω, ω') è lineare rispetto a ciascuna delle variabili.

Essendo la moltiplicazione esterna associativa, ha senso il prodotto esterno di un numero di covettori maggiore di 2, assegnati in un certo ordine. In particolare, per ogni k-pla di indici $\tau = (\tau_1,\dots,\tau_k)$ ha senso il prodotto esterno dei k-covettori $dx_{\tau_1},\dots,dx_{\tau_k}$ di grado 1, e risulta:

$$dx_{\tau_1} \wedge dx_{\tau_2} \wedge \cdots \wedge dx_{\tau_k} = dx_{\tau_1,\tau_2,\dots,\tau_k} = dx_\tau$$

sicchè nella notazione (3) precedentemente introdotta il simbolo $\wedge$ rappresenta la moltiplicazione esterna.

Analogamente a quanto si è fatto per i covettori, si costruisce una base per lo spazio $\overset{+}{\Omega}{}_k$ dei k-vettori.

Per ogni k-pla $\tau = (\tau_1,\dots,\tau_k)$ di indici, ossia di elementi dell'insieme $\{1,2,\dots,n\}$, con ciascuno dei simboli:

$$dx^{\tau_1} \wedge dx^{\tau_2} \wedge \dots \wedge dx^{\tau_k} \quad , \quad dx^{\tau_1,\tau_2,\dots,\tau_k} \quad , \quad dx^\tau$$

si denota il k-vettore che ad ogni matrice

$$A = \begin{bmatrix} a_1 \\ \vdots \\ a_k \end{bmatrix} \quad ,$$

ad n colonne, associa il determinante della matrice $k \times k$ individuata dalle colonne $\tau_1,\dots,\tau_k$ di A (considerate nell'ordine). In particolare, se $\tau = i \in \{1,\dots,n\}$ l'1-vettore $dx^\tau = dx^i$ associa ad ogni matrice $[a]$ la componente i-ma del vettore $a \in R^n$: sostanzialmente dx^i coincide con il covettore dx_i.

Ogni κ-vettore $\overset{+}{\omega}$ e' una combinazione lineare dei κ-vettori dx^τ con $\tau \in I_\kappa$; denotando con $\overset{+}{\omega}^\tau$ il coefficiente di dx^τ si ha dunque :

$$\overset{+}{\omega} = \sum_{\tau \in I_\kappa} \overset{+}{\omega}^\tau dx^\tau .$$

Di conseguenza il corrispondente mediante $\overset{+}{\omega}$ olella matrice A dianzi considerata e' una combinazione lineare dei determinanti minori di ordine massimo estratti da A .

La definizione del prodotto scalare in $\overset{+}{\Omega}_\kappa$ e' del tutto analoga a quella formulata per Ω_κ . Lo stesso dicasi per la moltiplicazione esterna nell'insieme costituito da tutti i κ-vettori con $\kappa \in N$.

Nello spazio prodotto $\Omega_\kappa \times \overset{+}{\Omega}_\kappa$ si definisce l'applicazione di dualita' $\langle \cdot, \cdot \rangle$, ponendo :

$$\langle \omega, \overset{+}{\omega} \rangle = \sum_{\tau \in I_\kappa} \omega_\tau \overset{+}{\omega}^\tau \qquad \forall \omega = \sum_{\tau \in I_\kappa} \omega_\tau dx_\tau \ \text{ e } \ \forall \overset{+}{\omega} = \sum_{\tau \in I_\kappa} \overset{+}{\omega}^\tau dx^\tau .$$

Soffermiamoci ora sul caso $\kappa = 1$.

Avendo identificato il vettore colonna $[v]$, ad n righe, con la n-pla dei suoi elementi, i covettori di grado 1 si identificano con le forme lineari in R^n, e quindi lo spazio Ω_κ con $\kappa = 1$ si viene ad identificare con il duale di R^n (di conseguenza con R^n) .

Analogamente, avendo identificato $\overset{+}{R}^n$ con lo spazio dei vettori riga ad n colonne , si conclude che lo spazio $\overset{+}{\Omega}_\kappa$ con $\kappa = 1$ si viene ad identificare con R^n.

Con le predette identificazioni il numero $\langle \omega, \overset{+}{\omega} \rangle$ (dualita' tra l'1-covettore ω e l'1-vettore $\overset{+}{\omega}$) coincide con il prodotto scalare dei vettori $(\omega_1, \ldots, \omega_n)$, $(\overset{+}{\omega}_1, \ldots, \overset{+}{\omega}_n)$ olello spazio R^n.

ESEMPI

Parlando di "funzione di κ variabili" intendiamo qui una funzione $\omega : (v_1, \ldots, v_\kappa) \to R$ con $v_1 \in R^n, \ldots, v_\kappa \in R^n$. Esamineremo alcuni valori di $\kappa \leq n$, dato che i κ-covettori con $\kappa > n$ sono nulli .

Es.1 . Covettori di grado 1 (forme lineari in R^n).

Un 1-covettore e' una funzione di una sola variabile, del tipo:

$$\omega = \sum_{\tau_1 = 1}^{n} \omega_{\tau_1} dx_{\tau_1} \tag{4}$$

che ad ogni $v = \begin{bmatrix} t_1 \\ \vdots \\ t_n \end{bmatrix}$ associa il numero $\sum_{\tau_1 = 1}^{n} \omega_{\tau_1} t_{\tau_1}$.

Es.2. Covettori di grado 2 (forme bilineari in R^n).

Un 2-covettore e' una funzione di due variabili, del tipo:

(5) $$\omega = \sum_{r_1,r_2}^{1\dots n} \omega_{r_1,r_2}\, dx_{r_1} \wedge dx_{r_2} ,$$

che ad ogni matrice:

$$[v_1, v_2] = \begin{bmatrix} t_{11} & t_{12} \\ \cdots & \cdots \\ \cdots & \cdots \\ t_{n1} & t_{n2} \end{bmatrix}$$

associa il numero:

$$\sum_{r_1,r_2}^{1\dots n} \omega_{r_1,r_2} \det \begin{bmatrix} t_{r_1,1} & t_{r_1,2} \\ t_{r_2,1} & t_{r_2,2} \end{bmatrix} .$$

Ad esempio per $n=3$, alla matrice

$$\begin{bmatrix} t_{11} & t_{12} \\ t_{21} & t_{22} \\ t_{31} & t_{32} \end{bmatrix}$$

associa il numero:

$$\omega'_{1,2} \det \begin{bmatrix} t_{11} & t_{12} \\ t_{21} & t_{22} \end{bmatrix} + \omega'_{1,3} \det \begin{bmatrix} t_{11} & t_{12} \\ t_{31} & t_{32} \end{bmatrix} + \omega'_{2,3} \det \begin{bmatrix} t_{21} & t_{22} \\ t_{31} & t_{32} \end{bmatrix} ,$$

avendo posto $\omega'_{1,2} = \omega_{12} - \omega_{21}$, $\omega'_{13} = \omega_{1,3} - \omega_{31}$, $\omega'_{2,3} = \omega_{2,3} - \omega_{3,2}$.

Per $n=2$ il covettore (5) alla matrice

$$A = \begin{bmatrix} t_{11} & t_{12} \\ t_{21} & t_{22} \end{bmatrix}$$

associa il numero $(\omega_{1,2} - \omega_{2,1}) \det A$.

Es.3. Covettori di grado 3.

Un 3-covettore e' una funzione di tre variabili, del tipo:

(6) $$\omega = \sum_{r \in I_3} \omega_r\, dx_r = \sum_{(r_1,r_2,r_3) \in I_3} \omega_{r_1,r_2,r_3}\, dx_{r_1} \wedge dx_{r_2} \wedge dx_{r_3} ,$$

che ad ogni matrice:

$$[v_1, v_2, v_3] = \begin{bmatrix} t_{11} & t_{12} & t_{13} \\ \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots \\ t_{n1} & t_{n2} & t_{n3} \end{bmatrix}$$

associa il numero:

$$\sum_{(r_1,r_2,r_3) \in I_3} \omega_{r_1,r_2,r_3} \det \begin{bmatrix} t_{r_1,1} & t_{r_1,2} & t_{r_1,3} \\ t_{r_2,1} & t_{r_2,2} & t_{r_2,3} \\ t_{r_3,1} & t_{r_3,2} & t_{r_3,3} \end{bmatrix} .$$

Consideriamo ora il covettore (5), che denotiamo con ω', chiamando poi r_2, r_3 i due indici della sommatoria che figura in (5). Detto ω il covettore (4), il prodotto esterno dei covettori ω ed ω' di gradi rispettivi 1 e 2, e' il covettore di grado 3:

$$\omega \wedge \omega' = \sum_{r_1,r_2,r_3}^{1\dots n} \omega_{r_1} \omega_{r_2 r_3}\, dx_{r_1} \wedge dx_{r_2} \wedge dx_{r_3} .$$

Es.4. Prodotto esterno di due covettori di grado 2.

Dati i covettori di grado 2 :

$$\omega' = \sum_{r_1,r_2}^{1\dots n} \omega'_{r_1 r_2}\, dx_{r_1} \wedge dx_{r_2} \quad , \quad \omega'' = \sum_{r_3,r_4}^{1\dots n} \omega''_{r_3 r_4}\, dx_{r_3} \wedge dx_{r_4}$$

il prodotto esterno di ω' per ω'' e' il covettore di grado 4 :

$$\omega' \wedge \omega'' = \sum_{r_1,r_2,r_3,r_4}^{1\dots n} \omega'_{r_1 r_2}\, \omega''_{r_3 r_4}\, dx_{r_1} \wedge dx_{r_2} \wedge dx_{r_3} \wedge dx_{r_4} ,$$

il quale ad ogni matrice:

$$[v_1, v_2, v_3, v_4] = \begin{bmatrix} t_{11} & t_{12} & t_{13} & t_{14} \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ t_{n1} & t_{n2} & t_{n3} & t_{n4} \end{bmatrix}$$

associa il numero:

$$\sum_{r_1,r_2,r_3,r_4}^{1\dots n} \omega'_{r_1 r_2}\, \omega''_{r_3 r_4}\, \det \begin{bmatrix} t_{r_1 1} & t_{r_1 2} & t_{r_1 3} & t_{r_1 4} \\ t_{r_2 1} & t_{r_2 2} & t_{r_2 3} & t_{r_2 4} \\ t_{r_3 1} & t_{r_3 2} & t_{r_3 3} & t_{r_3 4} \\ t_{r_4 1} & t_{r_4 2} & t_{r_4 3} & t_{r_4 4} \end{bmatrix}$$

* * *

Definiamo ora le <u>forme differenziali di grado k</u>, che in sostanza sono i k-covettori aventi per coefficiente delle funzioni, definite in un aperto di R^n (sicche' i k-covettori fin qui considerati sono forme differenziali di grado k a coefficienti costanti).

Sia A un aperto di R^n, e per ogni $r \in I_k$ sia ω_r una funzione reale definita in A. Orbene la corrispondenza $\omega : A \to \Omega_k$ che ad ogni $x \in A$ associa il k-covettore:

$$\omega(x) = \sum_{r \in I_k} \omega_r(x)\, dx_r \qquad (7)$$

si chiama una <u>forma differenziale di grado k</u>, <u>definita in A</u>; brevemente, si chiama una <u>k-forma</u>, <u>definita in A</u>. Per uniformita' di linguaggio si parla anche di <u>0-forma</u> definita in A, intendendo con tale locuzione una funzione reale definita in A.

Le funzioni ω_r si chiamano <u>i coefficienti</u> della k-forma ω; se essi sono continui [risp. di classe C^m, risp. costanti] in A, si dice che ω e' <u>continua</u> [risp. <u>di classe C^m</u>, risp. <u>costante</u>] <u>in A</u>.

I precedenti esempi di k-covettori sono delle k-forme a coefficienti costanti, e se tali coefficienti si pensano come funzioni definite in A, i predetti esempi sono k-forme definite in A.

L'insieme delle k-forme si considera munito, nel modo usuale, della struttura di

spazio vettoriale. L'operazione di moltiplicazione esterna, introdotta per i covettori, si estende in modo ovvio alle k-forme.

Allo scopo di definire il differenziale esterno della k-forma ω definita mediante la (7), supponiamo ω di classe C^1. Per ogni $\tau \in I_k$ ha senso il differenziale della funzione ω_τ:

$$(8) \qquad d\omega_\tau = \sum_{i=1}^{n} \frac{\partial \omega_\tau}{\partial x_i} dx_i \ ,$$

che è una forma differenziale di grado 1, ossia una 1-forma; orbene il prodotto esterno della (8) per la k-forma dx_τ, che è una (k+1)-forma, si chiama differenziale esterno di $\omega_\tau dx_\tau$, e si denota con $d(\omega_\tau dx_\tau)$:

$$(9) \qquad d(\omega_\tau dx_\tau) = d\omega_\tau \wedge dx_\tau \ .$$

Di conseguenza si chiama differenziale esterno della k-forma $\omega = \sum_{\tau \in I_k} \omega_\tau dx_\tau$, e si denota con $d\omega$, la (k+1)-forma:

$$d\omega = \sum_{\tau \in I_k} (d\omega_\tau) \wedge dx_\tau \ .$$

Per uniformità di linguaggio, il differenziale (8) della funzione ω_τ, ossia della 0-forma ω_τ, si chiama anche differenziale esterno di ω_τ.

L'operazione di differenziazione esterna è ovviamente lineare. Qualunque siano:

$$\omega = \sum_{\tau \in I_k} \omega_\tau dx_\tau \quad , \quad \eta = \sum_{\sigma \in I_h} \eta_\sigma dx_\sigma \ ,$$

di classe C^1, si ha poi:

$$(10) \qquad d(\omega \wedge \eta) = (d\omega) \wedge \eta + (-1)^{kh} \omega \wedge d\eta \ .$$

Inoltre, se la k-forma ω è di classe C^2, per ogni funzione $a \in C^1(A)$ sussiste la relazione:

$$(11) \qquad d(a d\omega) = da \wedge d\omega$$

che formalmente si ottiene dalla (8) quando alla k-forma costante dx_τ subentra la (k+1)-forma $d\omega$. Notiamo che per $a(x)$ costante dalla (11) consegue:

$$(12) \qquad d(d\omega) = 0 \ .$$

Dimostriamo la (10). Osserviamo dapprima che qualunque siano $\tau \in I_k$, $\sigma \in I_h$, per ogni funzione $a \in C^1(A)$ risulta:

$$(13) \qquad d(a dx_\tau \wedge dx_\sigma) = da \wedge dx_\tau \wedge dx_\sigma \ .$$

La (13) è banale se una delle coordinate della k-pla τ è uguale ad una delle coordinate della h-pla σ, perchè in tal caso ambo i membri sono nulli; in caso contrario la (13) consegue dalla (9), dato che esistono ovviamente $p \in I_{k+h}$ e un intero m tali che:

$$dx_\tau \wedge dx_\sigma = (-1)^m dx_p \ .$$

Per la (13) si ha:

$$d(\omega_\tau \eta_\sigma dx_\tau \wedge dx_\sigma) = d(\omega_\tau \eta_\sigma) \wedge dx_\tau \wedge dx_\sigma = (\eta_\sigma d\omega_\tau + \omega_\tau d\eta_\sigma) \wedge dx_\tau \wedge dx_\sigma =$$

$$= (d\omega_\tau \wedge dx_\tau) \wedge (\eta_\sigma dx_\sigma) + (\omega_\tau d\eta_\sigma) \wedge dx_\tau \wedge dx_\sigma = d(\omega_\tau dx_\tau) \wedge (\eta_\sigma dx_\sigma) + (-1)^{kh} \omega_\tau dx_\tau \wedge d\eta_\sigma \wedge dx_\sigma =$$

$$= d(\omega_\tau dx_\tau) \wedge (\eta_\sigma dx_\sigma) + (-1)^{kh} \omega_\tau dx_\tau \wedge d(\eta_\sigma dx_\sigma) \ .$$

La somma dei differenziali a primo membro per $\tau \in I_{n,h}$, $\sigma \in I_k$ è $d(\omega \wedge \eta)$; nell'ultimo membro, la sommatoria del primo addendo è $(d\omega)\wedge\eta$, quella del secondo addendo è $(-1)^{hh}\omega\wedge d\eta$.

Dimostriamo ora la (11), esaminando dapprima il caso che ω sia una 0-forma. In altri termini, preliminarmente dimostriamo che per $a, b \in C^1(A)$ risulta:

$$d(adb) = da \wedge db . \tag{14}$$

Intanto per la (9) si ha, con $j \in \{1,\dots,n\}$:

$$d\left(a\frac{\partial b}{\partial x_j}dx_j\right) = d\left(a\frac{\partial b}{\partial x_j}\right)\wedge dx_j = \frac{\partial b}{\partial x_j}\sum_{i=1}^{n}\frac{\partial a}{\partial x_i}dx_i\wedge dx_j + a\sum_{i=1}^{n}\frac{\partial^2 b}{\partial x_i \partial x_j}dx_i\wedge dx_j$$

e quindi, essendo: $d(adb) = d\left(a\sum_{j=1}^{n}\frac{\partial b}{\partial x_j}dx_j\right) = \sum_{j=1}^{n} d\left(a\frac{\partial b}{\partial x_j}dx_j\right)$,

risulta:

$$d(adb) = \sum_{j=1}^{n}\frac{\partial b}{\partial x_j}\sum_{i=1}^{n}\frac{\partial a}{\partial x_i}dx_i\wedge dx_j + a\sum_{i,j}^{1\dots n}\frac{\partial^2 b}{\partial x_i \partial x_j}dx_i\wedge dx_j .$$

Il secondo addendo a secondo membro è nullo: infatti per $i=j$ si ha $dx_i\wedge dx_j = 0$, mentre per $i\neq j$, in virtù del teorema di Schwarz, i due termini corrispondenti alle coppie $(i,j),(j,i)$ sono opposti. Dalla precedente uguaglianza consegue allora:

$$d(adb) = \left(\sum_{i=1}^{n}\frac{\partial a}{\partial x_i}dx_i\right)\wedge\left(\sum_{j=1}^{n}\frac{\partial b}{\partial x_j}dx_j\right)$$

ossia la (14).

Orbene, si ha:

$$a\,d\omega = a\sum_{\tau\in I_n} d\omega_\tau \wedge dx_\tau = \sum_{\tau\in I_n} a\,d\omega_\tau \wedge dx_\tau$$

e quindi per la (10), essendo $d(dx_\tau)=0$ [cfr. (9) con $\omega_\tau(x)=1$]:

$$d(a\,d\omega) = \sum_{\tau\in I_n} d\left[(a\,d\omega_\tau)\wedge dx_\tau\right] = \sum_{\tau\in I_n} d(a\,d\omega_\tau)\wedge dx_\tau .$$

Tenuto conto della (14) ne segue:

$$d(a\,d\omega) = \sum_{\tau\in I_n} da\wedge d\omega_\tau\wedge dx_\tau = da\wedge\sum_{\tau\in I_n} d\omega_\tau\wedge dx_\tau$$

donde la (11).

Una k-forma η, di classe C^1, si dice <u>chiusa</u> quando risulta:

$$d\eta = 0 ,$$

cioè quando il differenziale esterno di η è la funzione identicamente nulla in $(\mathbb{R}^n)^{k+1}$.

Si dice che una k-forma η, con $k\neq 0$, è <u>esatta</u>, oppure che η è un <u>differenziale esatto</u>, quando esiste una $(k-1)$-forma ω, di classe C^1, tale da risultare:

$$d\omega = \eta . \tag{17}$$

Si dice allora che la $(k-1)$-forma ω è una <u>primitiva</u> di η. In tale ipotesi, qualunque sia la k-forma costante ω_0, la k-forma $\omega+\omega_0$ è a sua volta una primitiva di η; ovviamente se l'aperto A di $\mathbb{R}^n$, insieme alla definizione dei coefficienti di η, è connesso, allora tutte le primitive di η sono del tipo $\omega+\omega_0$.

Affinché una k-forma η di classe C^1 sia esatta, <u>è necessario che sia chiusa</u>: infatti se η ammette una primitiva ω, questa è di classe C^2 e dalla (17) consegue, ricordando la (12), che $d\eta = d(d\omega) = 0$. È possibile dimostrare che la predetta condizione necessaria affinché η sia esatta <u>è anche sufficiente nell'ipotesi che l'aperto A sia stellato</u>.

Nel caso di una forma differenziale lineare ($k=1$):

$$\eta = \sum_{i=1}^{n} \eta_i \, dx_i ,$$

la condizione che η sia chiusa equivale al verificarsi in A delle uguaglianze (si veda il successivo esempio 1, formula (22)):

$$\frac{\partial \eta_i}{\partial x_j} = \frac{\partial \eta_j}{\partial x_i} \qquad (i,j = 1,\dots,n),$$

e l'ipotesi A stellato può essere sostituita con quella che l'aperto A sia <u>semplicemente connesso</u>. Così si chiama un aperto A quando ogni curva chiusa contenuta in A, che sia di classe C^2 (nel senso che sia il codominio di una funzione di classe C^2 in un intervallo $[a,b]$) è <u>omotopa ad un punto di A</u> [1].

Prima di riportare qualche esempio sulla differenziazione delle k-forme, introduciamo una ulteriore definizione.

Notiamo intanto che fin qui abbiamo considerato k-forme in R^n, ossia a valori nello spazio $\Omega_k(R^n)$: avendo fissato l'intero n una volta per tutte, per semplicità abbiamo adoperato la notazione Ω_k in luogo di $\Omega_k(R^n)$, e parlando di k-forme abbiamo sottinteso il riferimento ad R^n. Ora dobbiamo considerare anche delle k-forme in R^m (con m eventualmente diverso da n), sicché il riferimento predetto non potrà essere sottinteso; inoltre l'insieme I_k, che dipende da n, sarà ora indicato con $I_{k,n}$ e accanto ad esso sarà considerato l'insieme $I_{k,m}$.

Ciò premesso, qualunque siano $\varphi_1(y),\dots,\varphi_n(y)$ funzioni reali definite in un aperto U di R^m, e qualunque siano:

$$\tau = (\tau_1,\dots,\tau_k) \in I_{k,n} \quad , \qquad \sigma = (\sigma_1,\dots,\sigma_k) \in I_{k,m}$$

poniamo:

$$\frac{\partial \varphi_\tau}{\partial y_\sigma} = \frac{\partial(\varphi_{\tau_1},\dots,\varphi_{\tau_k})}{\partial(y_{\sigma_1},\dots,y_{\sigma_k})} \quad , \qquad d\varphi_\tau = \sum_{\sigma \in I_{k,m}} \frac{\partial \varphi_\tau}{\partial y_\sigma} dy_\sigma .$$

A partire dalla funzione $\varphi = (\varphi_1,\dots,\varphi_m)$ si ottiene in tal modo un insieme di

---

(1) Una curva γ contenuta in A, di classe C^2, si dice <u>omotopa ad un punto $x_0 \in A$</u> quando esiste una funzione $\varphi(i,t)$ a valori in A, definita per $i \in [0,1]$ e $t \in [a,b]$, <u>soddisfacente alle seguenti condizioni</u>:

1) φ è di classe C^2, e risulta $\varphi(i,a) = \varphi(i,b)$ per ogni $i \in [0,1]$.

2) Le funzioni f_0, f_1 definite ponendo:

$$f_0(t) = \varphi(0,t) \quad , \quad f_1(t) = \varphi(1,t) \qquad \forall t \in [a,b]$$

hanno per codominio rispettivamente la curva γ e l'insieme $\{x_0\}$.

Notiamo che quando esiste una funzione φ siffatta, se per ogni $i \in [0,1]$ si considera la funzione f_i definita ponendo $f_i(t) = \varphi(i,t)$ $\forall t \in [a,b]$, per $i \neq 1$ il codominio γ_i di f_i è una curva chiusa contenuta in A, che coincide con γ per $i = 0$. Chiamando le curve γ_i per $i \neq 0$ <u>deformazioni di γ</u>, si può dire allora che è possibile <u>deformare</u> la curva γ in modo che sia sempre contenuta in A e <u>si riduca</u> al punto x_0.

k-forme in R^m : l'insieme descritto dalla k-forma $d\varphi_\tau$ al variare di τ in $I_{k,n}$.

Orbene, data la k-forma ω in R^n, definita in un aperto A di R^n dalla relazione:

$$\omega(x) = \sum_{\tau \in I_{k,n}} \omega_\tau(x)\, dx_\tau \qquad \forall x = (x_1,..,x_m) \in A,$$

se risulta $\varphi(U) \subseteq A$ ha senso considerare la k-forma $\omega^\#$ in R^m definita in U ponendo:

$$\omega^\#(y) = \sum_{\tau \in I_{k,m}} \omega_\tau(\varphi(y))\, d\varphi_\tau(y) \qquad \forall y = (y_1,...,y_m) \in U.$$

La k-forma $\omega^\#$ viene solitamente chiamata <u>il pull-back di ω determinato da φ</u> : noi preferiamo dire che $\omega^\#$ e' la k-forma in R^m <u>generata da ω tramite φ.</u>

Sembra opportuno, anche se ovvio, rilevare che non si tratta della k-forma composta mediante φ ed ω ; quest'ultima, diciamola η , e' definita ponendo:

$$\eta(y) = \sum_{\tau \in I_{k,n}} \omega_\tau(\varphi(y))\, dy_\tau \qquad \forall y \in U.$$

La $\omega^\#$ dianzi introdotta interviene nel calcolo dell'integrale di una k-forma, esteso ad una k-varieta' orientata di R^n.

ESEMPI

<u>Es.5.</u> Il differenziale esterno della forma differenziale lineare in n variabili, cioe' della 1-forma , di classe C^1 :

$$\omega = \sum_{j=1}^{n} \omega_j\, dx_j \tag{18}$$

e' la 2-forma (forma differenziale bilineare in n variabili):

$$d\omega = \sum_{j=1}^{n} (d\omega_j) \wedge dx_j = \sum_{j=1}^{n} \sum_{i=1}^{n} \frac{\partial \omega_j}{\partial x_i}\, dx_i \wedge dx_j . \tag{19}$$

Si tratta della funzione che ad ogni matrice :

$$[v_1, v_2] = \begin{bmatrix} t_{11} & t_{12} \\ \cdots & \cdots \\ t_{n1} & t_{n2} \end{bmatrix} \tag{20}$$

associa il numero :

$$\sum_{\tau_1,\tau_2}^{1...n} \frac{\partial \omega_{\tau_2}}{\partial x_{\tau_1}} \det \begin{bmatrix} t_{\tau_1,1} & t_{\tau_1,2} \\ t_{\tau_2,1} & t_{\tau_2,2} \end{bmatrix} . \tag{21}$$

Detti τ,σ due interi appartenenti all'insieme $\{1,2,...,n\}$ con $\tau<\sigma$, sicche' la coppia (τ,σ) e' elemento dell'insieme $I_{2,n}$, nella (19) l'addendo che si ottiene per $i=\tau$ e $j=\sigma$ e' $\frac{\partial \omega_\sigma}{\partial x_\tau}\, dx_\tau \wedge dx_\sigma$, e quello che si ottiene per $i=\sigma$ e $j=\tau$ e' $\frac{\partial \omega_\tau}{\partial x_\sigma}\, dx_\sigma \wedge dx_\tau = -\frac{\partial \omega_\tau}{\partial x_\sigma}\, dx_\tau \wedge dx_\sigma$; di conseguenza

il coefficiente del termine $dx_{z_1} \wedge dx_{z_2}$ e' $\left(\frac{\partial \omega_{z_2}}{\partial x_{z_1}} - \frac{\partial \omega_{z_1}}{\partial x_{z_2}}\right)$. Pertanto la (19) si può scrivere:

$$(22) \qquad d\omega = \sum_{(z_1,z_2)\in I_{2,n}} \left(\frac{\partial \omega_{z_2}}{\partial x_{z_1}} - \frac{\partial \omega_{z_1}}{\partial x_{z_2}}\right) dx_{z_1} \wedge dx_{z_2} .$$

La forma differenziale (18) per $n=3$ solitamente si scrive, nel punto (x,y,z):

$$(23) \qquad \omega(x,y,z) = X(x,y,z)\,dx + Y(x,y,z)\,dy + Z(x,y,z)\,dz$$

e quindi il suo differenziale esterno nel punto (x,y,z) si scrive [cfr. (22) con $n=3$]:

$$d\omega(x,y,z) = \left(\frac{\partial Z}{\partial y} - \frac{\partial Y}{\partial z}\right) dy \wedge dz + \left(\frac{\partial X}{\partial z} - \frac{\partial Z}{\partial x}\right) dz \wedge dx + \left(\frac{\partial Y}{\partial x} - \frac{\partial X}{\partial y}\right) dx \wedge dy .$$

Naturalmente le forme bilineari $dy \wedge dz$, $dz \wedge dx$, $dx \wedge dy$ associano alla matrice (20) con $n=3$ rispettivamente i determinanti delle matrici:

$$\begin{bmatrix} t_{21} & t_{22} \\ t_{31} & t_{32} \end{bmatrix} , \quad \begin{bmatrix} t_{31} & t_{32} \\ t_{11} & t_{12} \end{bmatrix} , \quad \begin{bmatrix} t_{11} & t_{12} \\ t_{21} & t_{22} \end{bmatrix} .$$

Per $n=2$ la (18) solitamente si scrive, nel punto (x,y):

$$\omega(x,y) = X(x,y)\,dx + Y(x,y)\,dy$$

e quindi il suo differenziale esterno nel punto (x,y) si scrive [cfr. (22) con $n=2$]:

$$d\omega(x,y) = \left(\frac{\partial Y}{\partial x} - \frac{\partial X}{\partial y}\right) dx \wedge dy .$$

Naturalmente la forma bilineare $dx \wedge dy$ associa alla matrice (20) con $n=2$ il determinante della matrice:

$$\begin{bmatrix} t_{11} & t_{12} \\ t_{21} & t_{22} \end{bmatrix} .$$

Es. 6. Pull-back di una 1-forma.

Sia ω la 1-forma definita nell'aperto A di R^n dalla relazione:

$$(24) \qquad \omega(x) = \sum_{j=1}^{n} \omega_j(x)\,dx_j \qquad \forall x = (x_1,\dots,x_n) \in A ,$$

e sia φ una funzione definita nell'aperto U di R^m, a valori in R^n, di classe C^1.

Posto $\varphi = (\varphi_1,\dots,\varphi_n)$, se risulta $\varphi(U) \subseteq A$ il pull-back di ω determinato da φ e' la 1-forma definita in U ponendo:

$$\omega^{\#}(y) = \sum_{j=1}^{n} \omega_j(\varphi(y))\,d\varphi_j(y) = \sum_{j=1}^{n}\sum_{s=1}^{m} \omega_j(\varphi(y)) \frac{\partial \varphi_j}{\partial y_s}\,dy \qquad \forall y = (y_1,\dots,y_m) \in U .$$

Supponiamo $n=3$, $m=1$ e scriviamo la (24) nella forma (23). Se $\varphi(t) = (x(t),y(t),z(t))$ con $t \in]a,b[$, il pull-back di $\omega(x,y,z)$ determinato da φ e' la 1-forma in una variabile definita ponendo, per ogni $t \in]a,b[$:

$$\omega^{\#}(t) = \left[X(x(t),y(t),z(t))\,x'(t) + Y(x(t),y(t),z(t))\,y'(t) + Z(x(t),y(t),z(t))\,z'(t)\right] dt .$$

Es.7. Il differenziale esterno della forma differenziale bilineare in n variabili, cioè della 2-forma, di classe C^1:

$$\omega = \sum_{(i_1,i_2)\in I_{2,n}} \omega_{i_1,i_2}\, dx_{i_1}\wedge dx_{i_2}$$

è la 3-forma:

$$d\omega = \sum_{(i_1,i_2)\in I_{2,n}} (d\omega_{i_1,i_2})\wedge dx_{i_1}\wedge dx_{i_2} = \sum_{(i_1,i_2)\in I_{2,n}} \sum_{i=1}^{n} \frac{\partial \omega_{i_1,i_2}}{\partial x_i}\, dx_i\wedge dx_{i_1}\wedge dx_{i_2}\,.$$

Ad esempio per $n=4$, tenuto conto che le forme differenziali di grado 3:

$$dx_2\wedge dx_1\wedge dx_3\,,\quad dx_2\wedge dx_1\wedge dx_4\,,\quad dx_3\wedge dx_1\wedge dx_4\,,\quad dx_3\wedge dx_2\wedge dx_4$$

sono rispettivamente opposte alle seguenti:

$$dx_1\wedge dx_2\wedge dx_3\,,\quad dx_1\wedge dx_2\wedge dx_4\,,\quad dx_1\wedge dx_3\wedge dx_4\,,\quad dx_2\wedge dx_3\wedge dx_4$$

si ha:

$$d\omega = \left(\frac{\partial \omega_{2,3}}{\partial x_1} - \frac{\partial \omega_{1,3}}{\partial x_2} + \frac{\partial \omega_{1,2}}{\partial x_3}\right) dx_1\wedge dx_2\wedge dx_3 + \left(\frac{\partial \omega_{2,4}}{\partial x_1} - \frac{\partial \omega_{1,4}}{\partial x_2} + \frac{\partial \omega_{1,2}}{\partial x_4}\right) dx_1\wedge dx_2\wedge dx_4 +$$

$$+ \left(\frac{\partial \omega_{3,4}}{\partial x_1} - \frac{\partial \omega_{1,4}}{\partial x_3} + \frac{\partial \omega_{1,3}}{\partial x_4}\right) dx_1\wedge dx_3\wedge dx_4 + \left(\frac{\partial \omega_{3,4}}{\partial x_2} - \frac{\partial \omega_{2,4}}{\partial x_3} + \frac{\partial \omega_{2,3}}{\partial x_4}\right) dx_2\wedge dx_3\wedge dx_4\,.$$

Es.8. Pull-back di una 2-forma.

Sia ω la 2-forma definita nell'aperto A di R^n dalla relazione:

$$\omega(x) = \sum_{(i_1,i_2)\in I_{2,n}} \omega_{i_1,i_2}(x)\, dx_{i_1}\wedge dx_{i_2} \qquad \forall x=(x_1,\dots,x_n)\in A\,, \tag{25}$$

e sia φ una funzione definita nell'aperto U di R^m, a valori in R^n, di classe C^1.

Posto $\varphi=(\varphi_1,\dots,\varphi_n)$, se risulta $\varphi(U)\subseteq A$ il pull-back di ω determinato da φ è la 2-forma definita ponendo, per ogni $y=(y_1,\dots,y_m)\in U$:

$$\omega^{\#}(y) = \sum_{(i_1,i_2)\in I_{2,m}} \omega_{i_1,i_2}(\varphi(y))\, d\varphi_{i_1,i_2}(y) = \sum_{(i_1,i_2)\in I_{2,m}} \omega_{i_1,i_2}(\varphi(y)) \sum_{(s_1,s_2)\in I_{2,m}} \frac{\partial(\varphi_{i_1},\varphi_{i_2})}{\partial(y_{s_1},y_{s_2})}\, dy_{s_1}\wedge dy_{s_2}\,.$$

Per $n=3$ la forma differenziale (25) solitamente si scrive, nel punto (x,y,z):

$$\omega(x,y,z) = X(x,y,z)\,dy\wedge dz + Y(x,y,z)\,dz\wedge dx + Z(x,y,z)\,dx\wedge dy \quad (1)\,.$$

Se $\varphi(u,v)=(x(u,v),y(u,v),z(u,v))$ con $(u,v)\in B$, il pull-back di ω determinato da φ è la 2-forma definita in B ponendo, per $(u,v)\in B$:

$$\omega^{\#}(u,v) = \left[X(x(u,v),y(u,v),z(u,v))\frac{\partial(y,z)}{\partial(u,v)} + Y(x(u,v),y(u,v),z(u,v))\frac{\partial(z,x)}{\partial(u,v)} + Z(x(u,v),y(u,v),z(u,v))\frac{\partial(x,y)}{\partial(u,v)}\right] du\wedge dv\,.$$

(1) Solitamente in luogo dei prodotti esterni si scrivono i prodotti, in quanto dx, dy, dz vengono intesi come delle variabili reali; si ecco per ogni (x,y,z) la $\omega(x,y,z)$ viene intesa come un polinomio omogeneo di secondo grado in 3 variabili. Per tal motivo la ω viene chiamata forma differenziale quadratica (invece che bilineare).

INTEGRALE (SUPERFICIALE) DI UNA k-FORMA ESTESO AD UNA VARIETA' k-DIMENSIONALE ORIENTATA

Supponiamo di aver fissato una volta per tutte un intero $n \geq 2$, e un intero $k \geq 1$ minore di n. Premettiamo alcune definizioni.

Detto M un sottoinsieme di R^n, le tre ipotesi che seguono sono a due a due equivalenti. Per ogni $x_0 \in M$ esiste un intorno aperto I di x_0 tale che $M \cap I$ sia:

— il codominio di una funzione di classe C^1:

$$\varphi : U \longrightarrow R^n \quad \text{con } U \text{ aperto di } R^k \,, \tag{1}$$

invertibile e con la matrice jacobiana di rango k ;

— il grafico di una funzione di classe C^1:

$$f : U \longrightarrow R^{n-k} \text{ con } U \text{ aperto di } R^k \,;$$

— l'insieme degli zeri di una funzione di classe C^1:

$$F : I \longrightarrow R \,,$$

con la matrice jacobiana di rango $n-k$ [equivalentemente: di rango $n-k$ in x_0].

Quando si verifica una di tali ipotesi, l'insieme M si chiama una varietà k-dimensionale di R^n, brevemente una k-varietà. La funzione φ [risp. f, F] viene detta una parametrizzazione locale di M (relativa al punto x_0) ; quando occorrerà, noi aggiungeremo la precisazione : della quale M è localmente (intorno ad x_0) il codominio [risp. il grafico, l'insieme degli zeri]. Se accade che tutta la k-varietà M è il codominio di φ [risp. il grafico di f, l'insieme degli zeri di F], noi diremo che φ [risp. f, F] è una parametrizzazione globale di M, aggiungendo se occorre la predetta precisazione senza la parola "localmente".

Per ogni $x \in M$ sia $T_M(x)$ lo spazio vettoriale generato dai vettori tangenti in x a tutte le curve regolari passanti per x e contenute in M : tale spazio ha dimensione k, e si chiama lo spazio tangente alla k-varietà M nel punto x. Dette $\varphi_{u_1}, \dots, \varphi_{u_k}$ le derivate parziali della funzione (1), per ogni $x \in M \cap I$ il k-vettore di R^n:

$$\frac{\varphi_{u_1}(u) \wedge \varphi_{u_2}(u) \wedge \dots \wedge \varphi_{u_k}(u)}{\| \varphi_{u_1}(u) \wedge \varphi_{u_2}(u) \wedge \dots \wedge \varphi_{u_k}(u) \|} \qquad \text{con } u = \overset{\wedge}{\varphi}(x) \tag{2}$$

e' un orientamento di $T_M(x)$ [1]; la corrispondenza che ad ogni $x \in M \cap I$ associa il k-vettore (2) si chiama un orientamento di $M \cap I$, precisamente l'orientamento indotto da φ. Indicando con J_φ la matrice jacobiana di φ, sia $\hat{J}_\varphi$ la radice quadrata del determinante del prodotto della trasposta di $J_\varphi(u)$ per $J_\varphi(u)$ [per un noto teorema di Binet, tale determinante e' uguale alla somma dei quadrati dei determinanti minori di ordine k della matrice $J_\varphi(u)$]; per il denominatore della (2) vale la relazione:

$$(3) \qquad \| \varphi_{u_1}(u) \wedge \varphi_{u_2}(u) \wedge \dots \wedge \varphi_{u_k}(u) \| = \hat{J}_\varphi(u) \ .$$

Si verifica facilmente che il predetto orientamento di $M \cap I$ e' una funzione continua, e perciò la parte $M \cap I$ di M e' una k-varietà orientabile: così si chiama infatti una qualunque k-varietà M' quando e' possibile associare ad ogni $x \in M'$ un orientamento $\alpha(x)$ in modo che la funzione α sia continua.

Un sottoinsieme aperto A di R^k e' una particolare varietà k-dimensionale di R^n (le k-varietà sono sottoinsiemi di R^n, non necessariamente di R^k); la restrizione ad A dell'orientamento positivo di R^k [cfr. (1)] si chiama l'orientamento positivo della k-varietà A. Per uniformità di linguaggio, un aperto A di R^n si chiama una n-varietà, della quale si chiama orientamento positivo quello di R^n.

Qualunque sia $p \in N - \{1\}$ consideriamo ora un sottoinsieme compatto D di R^p la cui frontiera sia una varietà (p-1)-dimensionale di R^p; supponiamo che per ogni $x_0 \in \partial D$ esista una parametrizzazione locale F di ∂D, definita in un aperto I di R^p intorno di x_0, tale che non solo $I \cap \partial D$ coincida con l'insieme degli zeri di F, ma inoltre l'insieme dei punti di I in cui F assume valore negativo coincida con $I \cap \mathring{D}$. In tali ipotesi noi diremo che D e' un compatto regolare di R^p; il vettore di R^p:

$$\nu(x_0) = \frac{\operatorname{grad} F(x_0)}{|\operatorname{grad} F(x_0)|}$$

si chiama versore normale esterno al compatto regolare D nel punto x_0 [tale locuzione è dovuta al fatto che per $t > 0$ "abbastanza piccolo" il punto $x_0 + t\nu(x_0)$ non appartiene a D].

Si dimostra che la frontiera di un compatto regolare D di R^p e' orientabile. Conside-

---

(1) Qualunque sia la base $(h_1, \dots, h_p)$ del sottospazio V di R^n $(1 \le p \le n)$, il p-vettore di R^n:

$$\alpha = \frac{h_1 \wedge h_2 \wedge \dots \wedge h_p}{\| h_1 \wedge h_2 \wedge \dots \wedge h_p \|}$$

si chiama un orientamento di V, precisamente l'orientamento indotto dalla base considerata. Si dimostra che gli orientamenti possibili sono soltanto due, tra loro opposti.

Detta $(e_1, \dots, e_n)$ la base canonica di R^n, l'orientamento indotto da $(e_1, \dots, e_p)$ si chiama l'orientamento positivo di R^p.

riamo un orientamento δ di ∂D, e supponiamo che per ogni $y \in \partial D$ l'orientamento $\delta(y)$ dello spazio tangente a ∂D in y sia indotto dalla base $(h_1(y), \dots, h_{p-1}(y))$, che possiamo supporre costituita da vettori di R^p con la p-ma componente nulla; ovviamente la p-pla di vettori di R^p:

$$(4) \qquad (\nu(y), h_1(y), \dots, h_{p-1}(y))$$

è una base di R^p. Orbene, si dice che δ è l'orientamento positivo della frontiera di D quando per ogni $y \in \partial D$ l'orientamento di R^p indotto dalla base (4) coincide con l'orientamento positivo di R^p [cfr. nota (1) pag. 15].

Sulla nozione di compatto regolare di R^p, con $p=k$, si fonda quella di varietà regolare di R^m dotata di bordo. Sia M una k-varietà di R^m ($k>1$), ed esista una k-varietà M_1 contenente M che ammetta una parametrizzazione globale φ, di cui M_1 è il codominio, definita in un aperto U (di R^k). Orbene, se $M = \varphi(\mathring{D})$, essendo D un compatto regolare contenuto in U, si dice che M è una k-varietà regolare dotata di bordo; l'insieme $\varphi(\partial D)$ si chiama il bordo di M e si denota con ∂M: è possibile dimostrare che si tratta di una varietà di dimensione $(k-1)$. Le definizioni testè introdotte sono lecite, perchè si dimostra che esse sono indipendenti dalla k-varietà M_1 e dalla parametrizzazione globale φ di M_1.

La $(k-1)$-varietà ∂M è orientabile. Se α è un orientamento di M, uno dei due orientamenti di ∂M viene detto coerente con α: allo scopo di precisarlo, supponiamo che α sia l'orientamento di M indotto dalla parametrizzazione φ dianzi considerata (precisamente, dalla restrizione di φ a $\mathring{D}$).

Detto x_0 un punto di ∂M e posto $y_0 = \bar{\varphi}^{-1}(x_0) \in \partial D$, sia ψ una parametrizzazione locale di ∂D relativa ad y_0, definita in un aperto W di R^{k-1}, e si abbia:

$$\psi(W) = J \cap \partial D ,$$

essendo J un intorno aperto di y_0. Supponiamo che l'orientamento di $J \cap \partial D$ indotto da ψ sia quello positivo (cioè sia la restrizione a $J \cap \partial D$ dell'orientamento positivo di ∂D). È facile verificare che esiste un intorno aperto I di x_0 tale da risultare:

$$\varphi \circ \psi(W) = I \cap \partial M ,$$

e che $\varphi \circ \psi$ è una parametrizzazione locale di ∂M, relativa al punto x_0. Orbene, un orientamento β di ∂M si dice coerente con l'orientamento α di M se per ogni $x_0 \in \partial M$ la restrizione di β ad $I \cap \partial M$ coincide con l'orientamento di $I \cap \partial M$ indotto da $\varphi \circ \psi$.

L'intero k, per l'ipotesi posta all'inizio (pag.14), è minore di n : peraltro, in base a quanto convenuto, l'interno di un compatto regolare D di R^n è una n-varietà. In accordo con le precedenti definizioni, si conviene che $\mathring{D}$ sia una n-varietà regolare con bordo, e che il bordo di $\mathring{D}$ sia la frontiera di D. Evidentemente, considerando la frontiera di D come il bordo della n-varietà $\mathring{D}$, un orientamento di ∂D è coerente con un orientamento di $\mathring{D}$ quando entrambi sono l'orientamento positivo (cfr.pag.15-16) o quello opposto.

* * *

Siano: M una k-varietà di R^m, dotata di una parametrizzazione globale φ di cui è il codominio, ed M_0 un sottoinsieme di M.

Se il sottoinsieme $\vec{\varphi}(M_0)$ di R^k è misurabile, si dice che M_0 è <u>k-misurabile</u> e il numero [cfr.(3)]:

$$\sigma_k(M_0) = \int_{\vec{\varphi}(M_0)} \hat{J}_\varphi(u)\, du$$

si chiama <u>la misura k-dimensionale</u> di M_0. Supponiamo che M_0 sia k-misurabile.

Data una funzione reale f, continua in M_0, se la funzione continua:

$$f(\varphi(u))\,\hat{J}_\varphi(u) \tag{5}$$

è sommabile nell'insieme $\vec{\varphi}(M_0)$, si dice che f è <u>k-integrabile su M_0</u>; in tale ipotesi l'integrale esteso a $\vec{\varphi}(M_0)$ della (5) si chiama <u>integrale superficiale di f esteso ad M_0</u>, e si denota con:

$$\int_{M_0} f\, d\sigma_k \,. \tag{6}$$

Eliminando l'ipotesi che M sia dotata di una parametrizzazione globale, supponiamo che esistano un numero finito $M_1,\dots,M_p$ di k-varietà la cui unione è M, ciascuna dotata di parametrizzazione globale (tale ipotesi è in particolare verificata se la k-varietà M è un insieme compatto).

Se per ogni $i \in \{1,\dots,p\}$ l'insieme $M_0 \cap M_i$ è k-misurabile, e se f è ivi k-integrabile, si dice che f è <u>k-integrabile su M_0</u>. Allo scopo di definire in tale ipotesi l'integrale (6), consideriamo p funzioni continue definite in M:

$$\theta_1, \theta_2, \dots, \theta_p \,,$$

a valori nell'intervallo $[0,1]$, tali che la loro somma sia costante uguale ad 1 e per ogni $i \in \{1,\dots,p\}$ risulti $\theta_i = 0$ in $M - M_i$ [si dimostra che è possibile costruire tali funzioni, e si dice che esse costituiscono una <u>partizione dell'unità, associata agli insiemi $M_1,\dots,M_p$</u>].

Orbene, l'integrale superficiale di f esteso ad M_0 si definisce ponendo:

$$\int_{M_0} f\, d\sigma_k = \sum_{i=1}^{\infty} \int_{M_0 \cap M_i} \theta_i f\, d\sigma_k \,. \tag{7}$$

Le definizioni introdotte sono lecite, perché è possibile dimostrare l'indipendenza da cp della k-misurabilità di M_0 e dell'integrale (6), nonché l'indipendenza dell'integrale (7) sia dalle k-varietà M_i sia dalle funzioni θ_i. Parimenti la nozione di k-integrabilità è indipendente dalle k-varietà M_i.

Prima di passare a definire l'integrale di una k-forma, osserviamo che in base a quanto convenuto le definizioni precedenti si estendono in modo naturale al caso $k=n$. La misura n-dimensionale coincide con la misura di Lebesgue (la funzione identica è una parametrizzazione globale della n-varietà R^n); se f è una funzione continua in un compatto regolare D di R^n, l'integrale superficiale di f su D coincide con l'integrale di Lebesgue di f esteso a D (si dimostra che un compatto regolare è misurabile secondo Lebesgue). La frontiera di D è una (n-1)-varietà orientabile e compatta: perciò f è (n-1)-integrabile su ∂D ed ha senso l'integrale superficiale di f esteso a ∂D (si dimostra che ∂D è (n-1)-misurabile).

* * *

Sia M una k-varietà orientabile, unione di un numero finito di k-varietà ciascuna dotata di una parametrizzazione globale di cui è il codominio (in particolare, sia M una k-varietà orientabile e compatta). Detto α un orientamento di M (pag. 15), supponiamo che per ogni $x \in M$ il k-vettore $\alpha(x)$ di R^n (elemento dello spazio $\overset{\wedge}{\Omega}_k$, pag. 4) sia dato dalla relazione:

$$\alpha(x) = \sum_{\tau \in I_k} \alpha^\tau(x)\, dx^\tau \,.$$

Consideriamo quindi una forma differenziale ω di grado k, definita in un aperto A di R^n, e diciamo M_0 un sottoinsieme di M contenuto in A; per ogni $x \in A$ il k-covettore $\omega(x)$ (elemento dello spazio Ω_k) sia dato dalla relazione:

$$\omega(x) = \sum_{\tau \in I_k} \omega_\tau(x)\, dx_\tau \,.$$

Se ω è continua in M_0, essendo α continuo la funzione reale f definita ponendo:

$$f(x) = \langle \omega(x), \alpha(x) \rangle = \sum_{\tau \in I_k} \omega_\tau(x)\, \alpha^\tau(x) \quad \forall x \in M_0 \quad (1)$$

è continua in M_0.

---

(1) Il simbolo $\langle \cdot, \cdot \rangle$ denota l'applicazione di dualità in $\Omega_k \times \overset{\wedge}{\Omega}_k$ introdotta a pag. 5.

Orbene, se M_0 è k-misurabile e se f è k-integrabile su M_0, l'integrale superficiale [pag.17]

$$\int_{M_0} f \, d\sigma_k$$

si chiama <u>integrale della k-forma ω esteso ad M_0</u>, e si denota con

$$\int_{M_0,\alpha} \omega \, .$$

L'indicazione dell'orientamento α si omette quando non vi è possibilità di equivoci.

Se M_0 è contenuto nel codominio di una parametrizzazione locale φ di M, e se l'orientamento α di M coincide con quello indotto da φ [1], si ha:

$$(8) \qquad \int_{M_0} \omega = \int_{\bar{\varphi}(M_0)} \omega^{\#}$$

sempreché esistano entrambi gli integrali. Nella (8) la k-forma $\omega^{\#}$ è il pull-back di ω determinato da φ [pag.11], e l'insieme $\bar{\varphi}^{-1}(M_0)$, che è una k-varietà in quanto sottoinsieme aperto di R^k, è orientato positivamente [cfr. pag.15].

<u>Teorema della divergenza. Se ω è una $(n-1)$-forma di classe C^1, definita in un compatto regolare D di R^n, si ha:</u>

$$(9) \qquad \int_D d\omega = \int_{+\partial D} \omega$$

<u>dove la frontiera di D è orientata positivamente, e l'orientamento di D è quello indotto dalla base canonica di R^n.</u>

Per chiarire il motivo del nome dato al teorema, osserviamo quanto segue.

Essendo ω una $(n-1)$-forma, il suo differenziale esterno è una n-forma. Per $k = n-1$ l'insieme I_k introdotto a pag. 3 è costituito da n elementi, diciamoli $z_1, \dots, z_n$ denotando con z_i la $(n-1)$-pla che si ottiene da $(1, \dots, n)$ eliminando l'intero i; supposto che sia:

$$\omega = \sum_{i=1}^{n} \omega_{z_i} \, dx_{z_i} \, ,$$

diciamo v il vettore $(v_1, \dots, v_n)$ di R^n le cui componenti sono date da $v_i = (-1)^{i-1} \omega_{z_i}$. Orbene, si verifica che risulta:

---

[1] Precisamente: detto M_1 il codominio di φ, supponiamo che la restrizione di α ad M_1 coincida con l'orientamento di M_1 indotto da φ [cfr. (2) pag.14].

$$d\omega = (\operatorname{div} v)\, dx_1 \wedge dx_2 \wedge \ldots \wedge dx_n ,$$

e che il primo membro della (9) è l'integrale superficiale esteso a D della divergenza di v. Quest'ultimo, essendo $D \subset R^n$ (cioè sottoinsieme della n-varietà R^n) si riduce all'integrale di Lebesgue della divergenza di v [si tenga presente quanto detto a pag.17 prima della definizione di integrale].

Quanto al secondo membro della (9), se per ogni $x \in \partial D$ si denota con $\nu(x)$ il versore normale esterno a ∂D nel punto x [cfr. pag.15], si verifica che risulta:

$$\int_{+\partial D} \omega = \int_{\partial D} (v, \nu)\, d\sigma_{n-1}$$

essendo (v,ν) il prodotto scalare dei due vettori.

Teorema di Stokes. Sia ω una $(k-1)$-forma di classe C^1 in un aperto A di R^n, e sia M una varietà regolare dotata di bordo, tale che $M \cup \partial M \subset A$.

Se M e ∂M sono orientati coerentemente, risulta:

$$\int_{\partial M} \omega = \int_M d\omega$$

sempreché i due integrali abbiano senso.

UNA FUNZIONE CONTINUA IN UN INTERVALLO, CON DERIVATA NULLA QUASI OVUNQUE, CHE NON E' COSTANTE .

L'esempio in questione si costruisce utilizzando il cosiddetto *insieme di Cantor*, che si ottiene effettuando una opportuna successione di partizioni di un intervallo, ad esempio dell'intervallo [0,1].

Per comodità di esposizione, in via preliminare conveniamo quanto segue:

a) Qualunque siano gli intervalli disgiunti I e J, per esprimere che ogni numero appartenente ad I è minore di ogni numero appartenente a J scriveremo $I<J$ oppure $J>I$.

b) Ogni qualvolta prenderemo in considerazione una partizione dell'intervallo [0,1] costituita da un numero finito di intervalli, dopo aver elencato tali intervalli secondo la relazione < introdotta in **a**) chiameremo **laterali** quelli di posto dispari, che supporremo chiusi, e **centrali** quelli di posto pari, che supporremo aperti.

Tali locuzioni derivano dal fatto che occorrerà più volte suddividere un intervallo chiuso in tre intervalli: uno "centrale" e due "laterali".

E' opportuno evidenziare che:

1. *Se I e J sono intervalli disgiunti , essi sono anche separati e quindi se non risulta* $I \le J$ *si ha necessariamente* $I > J$.

Ciò premesso, andiamo a definire induttivamente la successione di partizioni $(\mathbf{P}_n)_{n\in\mathbf{N}}$ dell'intervallo [0,1] .

Partizione $\mathbf{P_1}$ – Si ottiene suddividendo l'intervallo [0,1] in tre intervalli di uguale ampiezza.

Partizione $\mathbf{P}_n$ ($n\ge 2$) – Si ottiene dalla partizione $\mathbf{P}_{n-1}$ suddividendo ciascuno degli intervalli laterali in tre intervalli di uguale ampiezza ([1]).

Verifichiamo che:

2. *Il numero degli intervalli laterali della partizione* $\mathbf{P}_n$ *è* 2^n *, quello degli intervalli centrali è* $2^n - 1$.

Dimostriamo che è vera la proposizione enunciata, diciamola $\mathbf{a}_n$), in base al principio di induzione. Essendo vera la $\mathbf{a}_1$), mostriamo che se è vera la $\mathbf{a}_n$) di conseguenza è vera la $\mathbf{a}_{n+1}$).

Il numero degli intervalli laterali della partizione $\mathbf{P}_{n+1}$ è il doppio del numero di quelli della $\mathbf{P}_n$ (ciascuno dei quali ha dato luogo, con la suddivisione, a 2 intervalli laterali) e quindi, essendo per ipotesi vera la $\mathbf{a}_n$), tale numero è $2\cdot 2^n = 2^{n+1}$.

Gli intervalli centrali della partizione $\mathbf{P}_{n+1}$ sono di duplice provenienza: quelli della partizione $\mathbf{P}_n$, rimasti inalterati, il numero dei quali è $2^n - 1$, e quelli ottenuti dalla suddivisione dei 2^n intervalli laterali della $\mathbf{P}_n$ (ciascuno dei quali ha dato luogo, con la suddivisione, ad un intervallo centrale della $\mathbf{P}_{n+1}$) : pertanto il numero degli intervalli centrali della $\mathbf{P}_{n+1}$ è $(2^n - 1) + 2^n = 2\cdot 2^n - 1 = 2^{n+1} - 1$.

([1]) Gli intervalli centrali della $\mathbf{P}_{n-1}$ rimangono inalterati nella partizione $\mathbf{P}_n$ **.** Di conseguenza, per ogni intero positivo $h<n$ gli intervalli centrali della partizione $\mathbf{P}_h$ sono anche intervalli centrali della $\mathbf{P}_n$ **.**

Con riferimento alla partizione $\mathbf{P_1}$, detti L_1^1 , L_2^1 gli intervalli laterali e C_1^1 quello centrale, si ha:

$$L_1^1 < C_1^1 < L_2^1 \ .$$

Per ogni $n \geq 2$ denotiamo con $L_i^n \ \left(i = 1,...,2^n\right)$ gli intervalli laterali della partizione $\mathbf{P_n}$; quanto agli intervalli centrali, supposto di aver denotato quelli della partizione $\mathbf{P_{n-1}}$ con $C_j^{n-1} \ (j = 1,...,2^{n-1}-1)$, denotiamo quelli della $\mathbf{P_n}$ con $C_j^n \ (j = 1,...,2^n - 1)$, avendo posto :

$$C_{2r}^n = C_r^{n-1} \qquad \forall r \in \left\{1,...,2^{n-1}-1\right\} \qquad (^2) \ .$$

Evidentemente per gli intervalli della partizione $\mathbf{P_n}$ sussistono le relazioni:

$$\begin{array}{llllllll} & L_1^n & < & C_1^n & < & L_2^n & < & C_2^n & < \\ & L_3^n & < & C_3^n & < & L_4^n & < & C_4^n & < \\ & \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots & & & & & & & \\ (1) & L_{2k-1}^n & < & C_{2k-1}^n & < & L_{2k}^n & < & C_{2k}^n & < \\ & \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots & & & & & & & \\ & L_{2^n-3}^n & < & C_{2^n-3}^n & < & L_{2^n-2}^n & < & C_{2^n-2}^n & < \\ & L_{2^n-1}^n & < & C_{2^n-1}^n & < & L_{2^n}^n & & & \end{array}$$

nelle quali gli intervalli della quarta "colonna" sono quelli centrali della partizione $\mathbf{P_{n-1}}$, rinominati come intervalli della $\mathbf{P_n}$.

Osserviamo che dalle (1) emergono le relazioni:

(2) $$C_{h-1}^n < L_h^n < C_h^n \qquad (n \geq 2,\ h \in [2,\ 2^n - 1] \cap \mathbf{N}) \quad (^3),$$

ed inoltre che:

3. *Qualunque siano* C_h^n *e* C_k^n, *intervalli centrali della partizione* $\mathbf{P_n}$, *per x,y soddisfacenti alle relazioni* :

(3) $$x \in C_h^n \ , \ y \in C_k^n$$

si ha :

(4) $$x<y \ \Rightarrow \ C_h^n \ \leq \ C_k^n \ \Leftrightarrow \ h \leq k \ . \quad (^4)$$

(^{2}) Abbiamo così "rinominato" ciascun intervallo centrale della partizione $\mathbf{P_{n-1}}$ [cfr. nota (^{1}) pag.1] , raddoppiandone l'indice ed aumentando l'esponente di una unità: la nuova notazione lo rappresenta come intervallo centrale della partizione $\mathbf{P_n}$.

Perciò gli intervalli centrali della $\mathbf{P_n}$ con l'indice pari sono tutti i centrali della $\mathbf{P_{n-1}}$; invece ciascuno di quelli con l'indice dispari proviene dalla suddivisione di un laterale della $\mathbf{P_{n-1}}$ in tre intervalli di uguale ampiezza (e dei tre è quello "centrale").

(^{3}) Osservando nelle (1) i simboli delle "colonne" seconda, terza e quarta, nelle varie "righe" (esclusa l'ultima) si vede la (2) con h pari, h=2,4,…,2^n–2; osservando nelle (1) l'ultimo simbolo di ciascuna riga (esclusa l'ultima) e i primi due della riga successiva, si vede la (2) con h dispari, h=3,5,…,2^n–1.

(^{4}) Supposto $x<y$, per assurdo ammettiamo che non sia $C_h^n \leq C_k^n$: per la proposiz. **1** si ha allora $C_h^n > C_k^n$ e quindi, per il significato della notazione >, risulta $x>y$ contro l'ipotesi . L'equivalenza che figura in (4) è dovuta alle notazioni adottate [cfr.(1)].

Per i nostri scopi ha interesse l'insieme Ω unione degli intervalli centrali, con l'indice dispari, di tutte le partizioni della successione $(\mathbf{P}_n)_{n\in\mathbf{N}}$: tale sottoinsieme dell'intervallo [0,1] interviene nel teorema seguente .

4. *Per ogni* $m\in\mathbf{N}$ *l'unione degli intervalli centrali della partizione* $\mathbf{P}_m$:

(5) $$\bigcup_{h=1}^{2^m-1} C_h^m$$

è uguale all'unione di tutti quelli delle partizioni $\mathbf{P_1}, \mathbf{P_2}, \ldots, \mathbf{P}_m$ *con l'indice dispari*:

(6) $$\bigcup_{n=1}^{m}\left(\bigcup_{k=1}^{2^{n-1}} C_{2k-1}^n\right)$$ ([5]) .

L'insieme Ω *che si ottiene considerando l'unione* (6) *per n da* 1 *a* $+\infty$:

(7) $$\Omega = \bigcup_{n\in N}\left(\bigcup_{k=1}^{2^{n-1}} C_{2k-1}^n\right)$$

è misurabile secondo Lebesgue ed ha misura 1 .

Consideriamo l'intervallo C_h^m che figura nell'unione (5). Se h è dispari, tale intervallo figura anche nell'unione (6) [dato che nella (6) l'indice della lettera C è dispari e l'esponente n deve assumere anche il valore m] , e quindi C_h^m è contenuto nell'insieme (6) ; alla stessa conclusione si perviene se h è pari dato che, posto $h = 2^r q$ con q dispari, risulta $C_h^m = C_q^{m-r}$ ([6]) : essendo q dispari e $m\text{-}r\in[1,m]$, l'intervallo C_q^{m-r} [ossia C_h^m] è contenuto nell'insieme (6).

Pertanto l'insieme (5) è contenuto nell'insieme (6).

Per dimostrare l'inclusione opposta poniamo:

(8) $$\Omega_n = \bigcup_{k=1}^{2^n-1} C_{2k-1}^n \quad ,$$

sicché l'insieme (6) è l'unione $\Omega_1 \cup \Omega_2 \cup \ldots..\cup \Omega_m$. L'insieme Ω_m , essendo l'unione di alcuni intervalli centrali della partizione $\mathbf{P}_m$, è contenuto nell'insieme (5); per ogni $n < m$ l'insieme Ω_n è l'unione di alcuni intervalli centrali della partizione $\mathbf{P}_n$, cioè di alcuni intervalli del tipo C_p^n: quest'ultimo, essendo anche intervallo centrale della partizione $\mathbf{P}_m$ in quanto $n<m$ [cfr. nota ([1]) pag.1] , è pure contenuto nell'insieme (5) e dopo ciò è chiaro che l'insieme (6) è contenuto nell'insieme (5).

La prima asserzione dell'enunciato è dunque acquisita. Al fine di dimostrare la seconda, poiché Ω è l'unione degli insiemi della successione $(\Omega_n)_{n\in N}$, in primo luogo proviamo che essi sono a due a due disgiunti.

Procedendo per assurdo, ammettiamo che esista un punto x appartenente all'insieme $\Omega_n \cap \Omega_i$ con $n\neq i$, sicché per la (8) x appartiene a due intervalli centrali con gli indici dispari rispettivamente delle partizioni $\mathbf{P}_n$ e $\mathbf{P}_i$, diciamoli C_{2h-1}^n e C_{2k-1}^i . Se ad esempio $i < n$, l'intervallo C_{2k-1}^i è anche un intervallo centrale della partizione $\mathbf{P}_n$ [cfr. nota ([1]) pag.1], e la notazione di esso come intervallo della $\mathbf{P}_n$ ha l'indice pari [cfr. nota ([2]) pag.2], quindi diverso da $2h$–1 : pertanto deve essere $i = n$. Dunque il punto x appartiene a due intervalli centrali della medesima partizione , il che è assurdo in quanto tali intervalli sono disgiunti.

([5]) Vogliamo chiarire il motivo per cui nell'unione tra parentesi l'indice k varia da 1 a 2^{n-1} . Si tratta dell'unione degli intervalli centrali della partizione $\mathbf{P}_n$ con l'indice dispari: siccome il numero di tutti quelli centrali è 2^n–1 [cfr. proposiz.**2** pag.1] , il numero di quelli che hanno l'indice dispari è [$(2^n-1)+1$] / 2 , ossia 2^{n-1} .

([6]) Prima di stabilire tale uguaglianza osserviamo che, essendo $2^r q = h < 2^m$, si ha $r<m$ [dato che $2^r \leq 2^r q < 2^m$] nonché $q<2^{m-r}$, quindi $q\leq 2^{m-r}$-1: perciò il simbolo C_q^{m-r} [introdotto a pag.2 prima delle (1)] ha significato, e rappresenta un intervallo centrale della partizione $\mathbf{P}_{m-r}$. Ricordando la nota ([2]) di pag.2, se vogliamo rappresentare C_q^{m-r} come intervallo della partizione $\mathbf{P}_m$ dobbiamo rinominarlo r volte : cioè aumentare l'esponente di una unità r volte (quindi sostituire m-r con m) e raddoppiare l'indice r volte (quindi moltiplicarlo per 2^r) . Perciò si ha $C_q^{m-r} = C_{2^r q}^m = C_h^m$.

Poiché ciascuno degli insiemi della successione $(\Omega_n)_{n\in N}$ è aperto in quanto unione di intervalli aperti, l'unione Ω di essi è a sua volta aperto e quindi è misurabile secondo Lebesgue. Siccome i predetti insiemi sono a due a due disgiunti si ha allora:

(9) $$\text{mis.}\,\Omega = \sum_{n\in N} mis.\Omega_n \ .$$

Allo scopo di determinare la misura di Ω_n, e quindi quella di Ω, mostriamo che per ogni $n\in\mathbf{N}$ gli intervalli laterali della partizione $\mathbf{P}_n$, e quelli centrali con l'indice dispari, hanno ampiezza $1/3^n$. Tale asserzione, diciamola $\mathbf{b}_n$, è vera per n=1 : procedendo per induzione, supposto n>1 ammettiamo che sia vera la $\mathbf{b}_{n-1}$ e proviamo che di conseguenza è vera la $\mathbf{b}_n$.

Gli intervalli centrali della partizione $\mathbf{P}_{n-1}$ **,** nelle notazioni che li rappresentano come intervalli della $\mathbf{P}_n$ hanno l'indice pari [cfr. nota ([2]) pag.2] e perciò non interessano ai nostri scopi. I rimanenti intervalli della $\mathbf{P}_n$, cioè quelli laterali e quelli centrali con l'indice dispari, sono stati ottenuti suddividendo in tre intervalli di uguale ampiezza gli intervalli laterali della $\mathbf{P}_{n-1}$, che per l'ipotesi induttiva hanno ampiezza $1/3^{n-1}$: pertanto essi hanno ampiezza $1/3^n$ e dunque la $\mathbf{b}_n$ è vera.

L'insieme Ω_n è l'unione dei 2^{n-1} intervalli centrali con l'indice dispari della partizione $\mathbf{P}_n$ [cfr. nota ([5]) pag.3]: siccome ciascuno di essi, per quanto dianzi stabilito, ha misura $1/3^n$, ne consegue che la misura di Ω_n è $2^{n-1}/3^n$. Dopo ciò dalla (9) si deduce:

$$\text{mis.}\,\Omega = \sum_{n\in N} mis.\,\Omega_n = \sum_{n\in N} 2^{n-1}/3^n = \frac{1}{3}\sum_{n\in N}\left(\frac{2}{3}\right)^{n-1} = 1$$

e l'asserto è dimostrato**.**

Orbene , il complemento rispetto all'intervallo [0,1] dell'insieme (7):

(10) $$K = [0,1] - \Omega$$

ovviamente chiuso [Cap.I pag.5], si chiama **l'insieme di Cantor** e per la proposizione **4** *ha misura nulla secondo Lebesgue*. Si dimostra, ma ciò non interessa per il seguito, che esso è equipotente all'intervallo [0,1] , sicchè è un insieme infinito non numerabile; vedremo tra breve che il suo complemento rispetto a [0,1], ossia Ω, è denso in [0,1].

* * *

L'insieme di Cantor consente di costruire una funzione, definita e continua nell'intervallo [0,1], che ha derivata nulla quasi ovunque e non è costante.

In via preliminare, diamo un cenno del procedimento. In primo luogo si definisce, nell'aperto Ω dato dalla (7), una funzione *f* costante a tratti, quindi con derivata nulla in ogni punto di Ω, e dalla definizione emerge che la *f* non è costante: poichè $[0,1]-\Omega$ è l'insieme di Cantor che ha misura nulla, ne segue che la *f* è definita *quasi ovunque* in [0,1] ed ha derivata nulla *quasi ovunque* in [0,1]. Successivamente si prova che *f* è crescente ; indi si dimostra che ogni punto dell'insieme di Cantor è di accumulazione per Ω ([7]), e che la funzione è ivi convergente: pertanto la *f* , continua in Ω perché derivabile, è prolungabile per continuità nei punti dell'insieme di Cantor. Il prolungamento di *f* , definito in [0,1], realizza la predetta costruzione.

Se *C* è un qualunque intervallo centrale di una qualunque delle partizioni $\mathbf{P}_n$, per semplicità di esposizione il valore comune che sarà attribuito ad *f* nei punti di *C* sarà indicato, sia pure impropriamente, con *f* (*C*). Orbene, ponendo:

([7]) Di conseguenza ogni punto dell'insieme di Cantor $[0,1]-\Omega$ è aderente ad Ω , e siccome anche i punti di Ω sono aderenti ad Ω , ne segue che $\overline{\Omega} = [0,1]$: dunque l'insieme Ω è denso nell'intervallo [0,1] (Cap.I pag.19).

(11) $$f\left(C_h^m\right)=\frac{h}{2^m} \qquad \forall\, m\in \mathbf{N} \quad \text{e} \quad \forall\, h\in[1,2^m-1]\cap \mathbf{N},$$

si definisce una funzione (denotata appunto con *f*) in tutti gli intervalli centrali della partizione $\mathbf{P}_m$, quindi nell'insieme (5) di pag.3, che per la proposizione **4** coincide con l'insieme (6). Per l'arbitrarietà di *m* ne consegue che *la funzione f è definita nell'insieme* Ω dato dalla (7) ([8]) .

Siccome a causa delle notazioni adottate può accadere che risulti

(12) $$C_h^m = C_k^n \quad \text{con } (m,h)\neq(n,k),$$

affinchè la definizione (11) sia lecita occorre stabilire che dalla (12) consegue:

(13) $$\frac{h}{2^m}=\frac{k}{2^n}\,.$$

A tale scopo osserviamo che i numeri naturali *h* e *k* si possono esprimere nella forma:

(14) $$h = 2^r p \quad , \quad k = 2^s q$$

con $p,q\in\mathbf{N}$ entrambi dispari e $r, s \in \mathbf{N_0}$, sicché risulta [cfr. nota ([6]) pag.3]:

$$C_p^{m-r} = C_{2^r p}^m = C_h^m \quad , \quad C_q^{n-s} = C_{2^s q}^n = C_k^n\,.$$

Dalla (12) consegue allora l'uguaglianza:

(15) $$C_p^{m-r} = C_q^{n-s} \quad ,$$

e questa comporta, come andiamo a dimostrare, che risulta:

(16) $$m\text{-}r = n\text{-}s \quad , \quad p = q \quad .$$

Se ad esempio fosse $m\text{-}r < n\text{-}s$, l'intervallo a primo membro della (15) sarebbe anche un intervallo della partizione $\mathbf{P}_{n\text{-}s}$ [nota ([1]) pag.1], e nella notazione che lo rappresenta come intervallo della $\mathbf{P}_{n\text{-}s}$ [quella a secondo membro della (15)] dovrebbe avere l'indice pari [nota ([6]) pag.3], in contrasto col fatto che *q* è dispari. Dunque sussiste la prima delle (16).

La seconda delle (16) è conseguenza della prima : i simboli al primo e secondo membro della (15) , avendo esponenti uguali, rappresentano un intervallo della medesima partizione, e quindi non possono avere indici diversi [basta osservare le (1) di pag.2].

Orbene, tenuto conto delle (14), (16) si ha:

$$\frac{h}{2^m}=\frac{2^r p}{2^m}=\frac{p}{2^{m-r}}=\frac{q}{2^{n-s}}=\frac{2^s q}{2^n}=\frac{k}{2^n}$$

e la (13) è dimostrata.

Dimostriamo ora che:

5. *La funzione f è crescente nel suo insieme di definizione* Ω .

Siano *x*,*y* punti di Ω , con

(17) $$x < y\,.$$

Poiché Ω è espresso dalla (7) di pag.3, tali punti appartengono a due intervalli centrali: per quanto detto nella nota ([1]) di pag.1 si può supporre che siano intervalli della stessa partizione . Dunque esistono $n,h,k\in\mathbf{N}$ tali che si abbia:

(18) $$x\in C_h^n \quad , \quad y\in C_k^n\,,$$

e quindi per la definizione (11), ricordando il significato del simbolo al primo membro, si ha :

$$f(x)=\frac{h}{2^n} \quad , \quad f(y)=\frac{k}{2^n} \quad .$$

Siccome le (17),(18) comportano che $h\le k$ [per la (4) di pag.2], dalle precedenti uguaglianze si deduce che $f(x)\le f(y)$.

Avendo acquisito che dalla (17) consegue $f(x)\le f(y)$, l'asserto è dimostrato.

([8]) Notiamo che risulta:

$$0 < f(x) < 1 \quad \forall\, x\in\Omega\,.$$

Per completare la costruzione della funzione dianzi descritta basta dimostrare che:

6. *Ciascun punto dell'insieme di Cantor è di accumulazione per* Ω *, e in esso la funzione* f *è convergente* : *di conseguenza* f *è prolungabile per continuità su tutto l'intervallo* [0,1].

Il prolungamento di f assume il valore 0 *nel punto* 0, *il valore* 1 *nel punto* 1 ([9]).

Fissato un qualunque punto x_0 dell'insieme di Cantor K , e posto :

(19) $$\Omega_0^- = \{x \in \Omega : x < x_0\} \quad . \quad \Omega_0^+ = \{x \in \Omega : x > x_0\} ,$$

dimostriamo intanto che se x_0 è interno all'intervallo [0,1] risulta :

(20) $$\sup \Omega_0^- = x_0 = \inf \Omega_0^+ ,$$

mentre se x_0=1 sussiste solo la prima delle (20) [dato che l'insieme Ω_0^+ è vuoto] , e se x_0=0 sussiste solo la seconda [dato che l'insieme Ω_0^- è vuoto] : ne seguirà in ogni caso che il punto x_0 è di accumulazione per Ω .

Supposto $x_0 \neq 0$ [ipotesi che comprende il caso x_0=1], il numero $a = \sup \Omega_0^-$ è ovviamente $\leq x_0$: se fosse $a < x_0$, l'intervallo $]a, x_0[$ sarebbe contenuto in K , il che è assurdo in quanto K ha misura nulla: dunque risulta $a = x_0$ e la prima delle (20) è acquisita. Analogamente , supposto $x_0 \neq 1$ [ipotesi che comprende il caso x_0=0], si conclude che sussiste la seconda delle (20).

Dimostriamo ora, in primo luogo, che f è prolungabile per continuità nei punti 0 , 1 e che vale l'ultima asserzione dell'enunciato.

Poiché la funzione f è crescente , essendo limitata inferiormente [0 è un minorante] essa è convergente nel punto 0, ed essendo limitata superiormente [1 è un maggiorante] è convergente nel punto 1; d'altra parte, se x_0=0 [x_0=1] si ha $\Omega_0^+ = \Omega$ [$\Omega_0^- = \Omega$] e quindi, per la (20), $\inf \Omega = 0$ [$\sup \Omega = 1$]. Pertanto, detto f^* il prolungamento per continuità di f su $\Omega \cup \{0,1\}$, risulta:

(21) $$f^*(0) = \lim_{x \to 0+} f(x) = \inf_{x \in \Omega} f(x) \quad , \quad f^*(1) = \lim_{x \to 1-} f(x) = \sup_{x \in \Omega} f(x) .$$

Scelto arbitrariamente $m \in \mathbf{N} - \{1\}$, siano ora x_m un punto di $C_1^m \subseteq \Omega$ e y_m un punto di $C_{2^m-1}^m \subseteq \Omega$. Tenuto conto della nota ([8]) di pag.5, dalle (21) consegue, ricordando la (11):

$$0 \leq f^*(0) = \inf_{x \in \Omega} f(x) \leq f(x_m) = \frac{1}{2^m}$$

$$1 \geq f^*(1) = \sup_{x \in \Omega} f(x) \geq f(y_m) = \frac{2^m - 1}{2^m} ,$$

sicché risulta:

$$0 \leq f^*(0) \leq \frac{1}{2^m} \quad , \quad \frac{2^m - 1}{2^m} \leq f^*(1) \leq 1 .$$

Per l'arbitrarietà di m, se ne deduce

$$f^*(0) = 0 \quad , \quad f^*(1) = 1 .$$

Resta da dimostrare che per ogni punto x_0 interno all'intervallo [0,1] , la f è convergente in x_0 . Intanto osserviamo che, essendo f crescente, esistono finiti i limiti sinistro e destro di f in x_0 ; posto:

(22) $$\lambda_1 = \lim_{x \to x_0-} f(x) = \sup_{x \in \Omega_0^-} f(x) \quad , \quad \lambda_2 = \lim_{x \to x_0+} f(x) = \inf_{x \in \Omega_0^+} f(x) ,$$

si ha ovviamente

(23) $$\lambda_1 \leq \lambda_2 .$$

Si tratta dunque di dimostrare che $\lambda_1 = \lambda_2$.

Scegliamo arbitrariamente $m \in \mathbf{N} - \{1\}$ e consideriamo la partizione $\mathbf{P}_m$. Siccome x_0 non appartiene ad Ω , per la proposizione **4** non appartiene ad alcuno degli intervalli centrali della $\mathbf{P}_m$ ([10]). Dunque il punto x_0 appartiene ad uno degli intervalli laterali della partizione $\mathbf{P}_m$; supponiamo allora che si abbia :

([9]) Gli estremi dell'intervallo [0,1] sono punti dell'insieme di Cantor : basta osservare che il punto 0 [il punto 1] appartiene al primo [all'ultimo] intervallo *laterale* di qualunque partizione della successione $(\mathbf{P}_n)_{n \in \mathbf{N}}$, sicché non appartiene ad Ω .

(24) $$x_0 \in L^m_{h_m} \qquad \text{con } h_m \in [1, 2^m] \cap \mathbf{N}.$$

Se h_m=1 qualunque sia $m \in \mathbf{N} - \{1\}$, siccome l'intervallo L^m_1 ha estremo sinistro 0 e ampiezza $1/3^m$ ([11]), risulta $0 \le x_0 \le 1/3^m$ e quindi $x_0 = 0$ **(** per l'arbitrarietà di m) : pertanto, avendo supposto $x_0 \neq 0$, esiste almeno un valore di m , diciamolo m_0 , tale che $h_{m_0} > 1$; allora anche per $m > m_0$ si ha $h_m > 1$ (basta osservare che per effettuare la partizione P_{m_0+1} si decompone $L^{m_0}_1$ in tre intervalli il primo dei quali è $L^{m_0+1}_1$: se fosse $h_{m_0+1} = 1$, cioè se il punto x_0 appartenesse all'intervallo $L^{m_0+1}_1$ **,** esso apparterrebbe anche all'intervallo $L^{m_0}_1$ che lo contiene e quindi sarebbe $h_{m_0} = 1$).

Con analogo discorso si conclude , avendo supposto $x_0 \neq 1$, che esiste un valore di m , diciamolo m_1 , tale che per $m > m_1$ risulti h_m minore del valore massimo possibile, cioè $h_m < 2^m$.

Orbene, d'ora in avanti supporremo m (sempre scelto arbitrariamente) maggiore sia di m_0 che di m_1 , sicchè l'intero h_m appartiene all'intervallo $[2,\ 2^m - 1]$. Possiamo allora utilizzare la (2) di pag.2 con $n=m$, $h=h_m$, cioè la relazione:

(25) $$C^m_{h_m-1} < L^m_{h_m} < C^m_{h_m} \ .$$

Scelti un punto $x_m \in C^m_{h_m-1}$ e un punto $y_m \in C^m_{h_m}$, per le (24), (25) si ha

(26) $$x_m < x_0 < y_m \ ,$$

e per la (11) risulta:

(27) $$f(x_m) = f(C^m_{h_m-1}) = \frac{h_m - 1}{2^m} \qquad , \qquad f(y_m) = f(C^m_{h_m}) = \frac{h_m}{2^m} \ .$$

D'altra parte la (26), ricordando la (19), comporta :

$$x_m \in \Omega_0^- \ , \quad y_m \in \Omega_0^+$$

e quindi per le (22), (23) si ha:

$$f(x_m) \le \lambda_1 \le \lambda_2 \le f(y_m) \ .$$

Se ne deduce che

$$0 \le \lambda_2 - \lambda_1 \le f(y_m) - f(x_m) \ ,$$

e quindi per le (27) risulta:

$$0 \le \lambda_2 - \lambda_1 \le \frac{1}{2^m} \ .$$

Per l'arbitrarietà di m ne segue che $\lambda_1 = \lambda_2$ e l'asserto è provato.

OSSERVAZIONE. Detto f^* il prolungamento per continuità di f su [0,1], la funzione f^* è una primitiva generalizzata [pag.47] della funzione identicamente nulla, diciamola ω , e non sussiste l'uguaglianza:

(28) $$\int_0^1 \omega(x)\, dx \ = \ f^*(1) \ - \ f^*(0),$$

in quanto il primo membro è uguale a zero mentre il secondo membro, per la proposizione **6,** è uguale ad 1. Ciò prova che : 1). La funzione f^* non è assolutamente continua [se fosse assolutamente continua dovrebbe valere la (28), come si è detto a pag.47] ; 2). Per la validità della formula fondamentale del calcolo integrale è necessario utilizzare una primitiva generalizzata che sia assolutamente continua.

([10]) L'insieme Ω è definito dalla (7) di pag.3 : il punto x_0 , non appartenendo ad Ω , qualunque sia m non appartiene all'insieme (6) , che coincide con l'insieme (5) [per la proposiz. **4**].

([11]) L'ampiezza degli intervalli laterali della partizione $\mathbf{P}_n$ è stata determinata a pag. 4 **,** dopo la (9) .

INDICE ANALITICO

A

aderenza 6
aggiunto : v. Operatore
applicazione:
  bilineare continua 209
  biunivoca [8]
  canonica di uno spazio
    [nel suo duale 187
  di dualità 318
  iniettiva [8]
  suriettiva [8]
assioma di Hausdorff 11
assiomi di Hausdorff 16
autospazio 240
autovalore:
  di una matrice 240, 251
  di un operatore lineare 240, 251
autovettore 240

B

base :
  di topologia 18
  di uno spazio vettoriale 50
biduale 186

C

cerchio, di uno spazio metrico 77
chiusura 6
  debole 200
classe:
  dei resti [4]
  di equivalenza [5], [2], [4]
coefficienti di Fourier 127
combinazione lineare 50
complementare 4
complemento 4
  ortogonale 96, 109
completamento:
  di uno spazio metrico 158.1
  di uno spazio normato 158.1
componenti di un vettore :
    di uno spazio di Hilbert 130
      a dimensione finita 119
cono di $\mathbf{R}^k$ 101 nota ([1])
continuità , vedi *funzione continua*
contrazione, di uno spazio metrico 285
convergenza:
  debole 107, 197
convergenza (*continuazione*) :
  debole degli operatori 208
  *debole (o puntuale) 191
  forte 197
    degli operatori 208
  in media 99
  in norma (di operatori) 208
  nello spazio quoziente 110
  puntuale :
    di operatori 206
    nel duale 191
condizione di Hölder 56
    di Lipschitz 56

D

decomposizione spettrale 247
densità, di una misura 45, 149
derivata:
  debole 59
  di una funzione a valori in
    [uno spazio pre-hilbertiano 117
  di una distribuzione 158.2
  nel senso delle distribuzioni 158.2
    della funzione scalino 158.3
derivato 7
  debole 200
diametro 79
differenziale 257.1 , 288
  di ordine superiore 298
  di Frechet 290
  di Gateau 257.1, 288
  in senso
    debole 257.1,288
    forte 290
dimensione, di uno spazio vettoriale 50
disequazione variazionale 116, 254,
    [257.2, 312, 317
  esempi di 257.4, 263, 271, 279
distanza 77, 79
distribuzione 151:
    di Dirac 154,157
    positiva 151, 153
    su un aperto Ω 151
disuguaglianza:
  di Bessel 128
  di Friedrics 104
  di Schwarz 41 nota ([1]), 94
  di Schwarz-Hölder 54
dominio, di un operatore 62
duale:
  algebrico 69

duale (*continuazione*):
dei seguenti spazi:
$C^0(\Omega)$ 182.1
$C_0^\infty(\Omega)$ 182.1
$H^{m,p}(\Omega)$ 182
$H_0^{m,p}(\Omega)$ 182.1
$L^p(\Omega)$ 70, 180
$L^1(\Omega)$ 182.1
$L^\infty(\Omega)$ 182
$\mathbf{R}^k$ 70, 179
$W^{m,p}(\Omega)$ 182.1
di uno spazio :
a dimensione finita 70, 180
di Hilbert 177,179
forte 173
topologico 171
dualità 69, 172

E

equazione:
di Eulero-Lagrange 298
differenziale in senso debole 158.5
di Laplace 182.4
di Poisson 182.4
variazionale 117, 257, 257.3
esempi di 277
equazioni integrali di Fredholm 168, 236, B.2
con nucleo di classe $N^{(\alpha)}$ B.13
introduzione alle A.1
estremi, di un operatore autoaggiunto 223
estremo:
inferiore essenziale 37
inferiore [superiore] di una funzione
[definita q.o. 42
superiore essenziale 36

F

famiglia :
di elementi di un insieme 19
ortogonale 122
ortonormale 122
finita additività 44
forma:
bilineare:
coercitiva 213
esempi di 213
hermitiana 67, 68
simmetrica 67, 68
differenziale di grado *k* C.1
forma (*continuazione*)
quadratica 67, 68 formula (7)
sesquilineare 67, 68 formula (6)
frontiera 7
funzionale 63
bilineare 67
complesso 63
convesso 63
lineare 63
strettamente convesso 63
funzionali lineari (esempi)
in spazi a dimensione finita 145,180
in spazi pre-hilbertiani 97, 177,179
nello spazio :
$C_0^k(\Omega)$ 66
$C_0^\infty(\Omega)$ 151
$C^0(\Omega)$ 182.1
$C^2(T)$ 66, 146
$C_0^\infty(\Omega)$ 182.1
$H^{m,p}(\Omega)$ 182
$H_0^1(\Omega)$ 148
$H_0^2(\Omega)$ 147, 148
$H_0^1(\Omega) \cap H^2(\Omega)$ 148
$H_0^m(\Omega)$ 147
$L^p(\Omega)$ 66, 146,180
$L^\infty(\Omega)$ 182
$\mathbf{R}$ 64
$\mathbf{R}^k$ 64, 144, 145, 179
$W^{m,p}(\Omega)$ 66, 147
funzione:
a variazione limitata 46
assolutamente continua 47, 330
rispetto alla misura di
[Lebesgue 45
continua 21, 43,110, 133,
[221 nota (1)
con derivata q.o. nulla D.1
quasi ovunque
[35 nota(2), 329
per successioni 109,151
convergente in un punto 20
definita quasi ovunque 42 :
continua 43
limitata 42
misurabile 42
sommabile 42
derivabile (a valori in uno spazio
[pre-hilbertiano) 117 nota (1)

funzione (*continuazione*)
differenziabile:
in senso
debole 257.1, 288
forte 290
secondo
Frechet 290
Gateau 257.1 ,288
di intervallo 48.2
di quadrato sommabile 41
essenzialmente limitata 36
hölderiana 56, 330
inversa 8
invertibile 8
integrabile secondo Riemann
[35 nota ([1])
limitata 42
misurabile 34,42
nel senso delle distribuzioni 152
numerabilmente additiva 44
peso 46
quasi continua 35, 329
scalino 158.1
sommabile:
rispetto
ad una funzione peso 46
ad una misura 45
secondo Lebesgue 40, 42
secondo Riemann 41
uniformemente continua 135, 330

G

gradiente 289
grafico 142
gruppo 49
abeliano 49
commutativo 49

I

identità : v. Operatore identico
immersione 63
di $C^{0,\lambda}(T)$ in $C^{0}(T)$ 170
di $C^{m,\lambda}(T)$ in $C^{m}(T)$ 170
di $C^{m+1}(T)$ in $C^{m}(T)$ 170
di $H^{m}(\Omega)$ in $C^{m-1,\,1/2}(\Omega)$ 105
di $W^{m,p}(\Omega)$ in vari spazi 304
iniezione 63
insieme :
aperto 4, 5
aperto a frontiera localmente
[lipschitziana 304
insieme (*continuazione*)
aperto in , su 9
chiuso 5, 6, 7
chiuso in , su 9
compatto 14
connesso 12, 86
convesso 52, 84
debolmente:
chiuso 200
sequenzialmente
compatto 200
precompatto 200
degli indici 19
delle potenze degli insiemi 15
denso 19
di Cantor D.4
discreto 18
finito 13
infinito 13
limitato 79
misurabile secondo Lebesgue 32, 33
numerabile 16 , 18
ortogonale 122
ortonormale 122
precompatto 15
relativamente compatto 15
risolvente 239
insiemi equipotenti 13 , 17
integrale:
del calcolo delle variazioni 296
di una funzione continua a valori
[in uno spazio di Banach 301
di Stieltjes 48, 48.1, 48.2, 306 nota ([3])
di Stieltjes-Lebesgue 45,149
secondo Lebesgue 39, 41, 43
rispetto:
ad una funzione peso 46
ad una misura 46
interno, di un insieme 6, 8
intorno 5
in, su 9
inverso di un operatore lineare 73
destro, sinistro 75
inviluppo convesso 52, 85
iperpiano 52, 63, 175
ipotesi del continuo 19
isomorfismo :
algebrico 62
canonico:
di $\mathbf{R}^k$ su uno spazio normato
[di dimensione n 141
di uno spazio di Hilbert sul duale 179
di uno spazio di Banach
[sul biduale 187
isometrico 140
di uno spazio di Hilbert
separabile su l_2 130
sul duale 179
tra spazi normati 139, 175, 229

L

laplaciano
  nel senso delle distribuzioni 182.2
lavoro di una forza 274
leggi di De Morgan 5
lemma di Minty 257.4
limite:
  debole 197
  di una funzione 20
  di una somma 86
  di una successione 19
  puntuale nel duale 191
  puntuale debole [forte] di una
    [successione di operatori 206

M

metodo delle approssimazioni
    [successive 284
metrica 77
metriche equivalenti 79
misura 44
    di Dirac 153
    dotata di densità 45,149
    secondo Lebesgue 32, 33
    spettrale 306
      associata ad un operatore 308
modelli matematici 182.4
molteplicità di un autovalore 240
    algebrica 251
multiindice 56

N

norma 88
    di un'applicazione bilineare continua 210
    di una forma sesquilineare continua 212
    di un funzionale lineare continuo 173
    di un multiindice 56
    di un operatore integrale 203bis
    di un operatore lineare
      [continuo 202, 203, 203bis, 208
    di un proiettore 203 bis
    in uno spazio a dimensione finita
      [203bis nota ([1])
    nel duale di uno spazio normato 173
norme equivalenti 92. 203
nucleo:
  dell'operatore integrale 169, B.2, B.11
  di un operatore lineare 62
  iterato B.2
  risolvente B.11
numerabile additività 34, 44
numero cardinale,
    [di un insieme 13

O

operatore:
  aggiunto 219, 227.5
    hermitiano 219, 232
    esempi di 227
  autoaggiunto 220, 222bis, 223
  maggiore di un altro 223
  positivo 223
  chiuso 142
  coercivo 257.2
  compatto 165
    esempi di 169, 170, 201
  completamente continuo 166
  di Laplace ,vedi *laplaciano*
  emicontinuo 257.2
  identico 63, 170 n.3
  integrale B.1, 169, 201,
    [203bis, 218, 227.4
  limitato 137
  lineare 62
    continuo 138, 142, 202
    inverso 73
    invertibile 73
  monotono 71
    strettamente 71
  unitario 311
operatori lineari (esempi)
    di $C^1(\Omega)$ in $C^0(\Omega)$ 203bis
    di $C^2(\Omega)$ in $C^0(\Omega)$ 66, 146
    di $H_0^2(\Omega)$ in $L^2(\Omega)$ 227.2, 227.3
    di $H_0^m(\Omega)$ in $L^2(\Omega)$ 147,
      [203bis, 227.1
    di $L^p(\Omega)$ in $L^p(\Omega)$ 66
    di $\mathbf{R}^k$ in $\mathbf{R}^m$ 64, 145
    tra spazi a dimensione finita 64, 145
ortogonali , con riferimento a :
    vettori di uno spazio pre-hilbertiano 96
    vettori il primo in uno
      [spazio pre-hilbertiano, il
        [secondo nel duale 174

P

parte:
  libera 50
  staccata 12
perturbazione, di una disequazione
    [variazionale 312
polinomi di:
    Hermite 132
    Laguerre 132
    Legendre 132
    Tchebicev 132

potenza:
del continuo 18
del numerabile 16
di un insieme 14 , 12 , 13
primitiva
debole 182.5
di un funzionale lineare e
[continuo definito in 158.4
generalizzata 47
principio:
delle contrazioni 285
di uniforme limitatezza 206
problema:
ben posto 182.4
omogeneo di Dirichlet
generalizzato 182.2
in forma debole 182.2,182.3

sul filo flessibile con ostacolo:
deformabile 267
rigido 260
sulla trave incastrata 274
con ostacolo 278

prodotto:
di operatori 72
di spazi :
di Banach 113
metrici 81
normati 93
topologici 17
vettoriali normati 93
di topologie 17
scalare 93, 94
in $\mathbf{R}^k$, $\mathbf{C}^k$ 98
in uno spazio a dim.finita 180
proiettore 121, 203 bis, 224
proiezione:
di un punto su un convesso, in uno
[spazio di Hilbert 115
ortogonale 118
proprietà :
di cono 101 nota ([1])
verificata quasi ovunque 36 nota ([1])
pseudodominio 56
punto:
aderente 6
di accumulazione 7
di uno spazio topologico 4
fisso 284
frontiera 7
interno 6
isolato 7

Q

quasi ovunque:
continua (riferito a *funzione*)
[35 nota ([2]), 329
definita (riferito a *funzione*) 42
verificata (riferito a
proprietà) [36 nota([1])

R

raggio spettrale 242
rappresentante 2 , 5
rappresentazione spettrale:
dell'identità 247, 308
di un operatore
autoaggiunto 308
compatto autoaggiunto 245
relazione:
binaria 1
di equivalenza 1
riflessiva 2
simmetrica 2
transitiva 2
relazioni di De Morgan 5
retta, di uno spazio vettoriale 52, 63

ricoprimento 6 , 13
aperto 14
finito 14
risolvente : v. Insieme, Nucleo

S

scalare 50, 93, 94
sfera, di uno spazio metrico 77
segmento, di uno spazio vettoriale 52
serie 125
di Fourier 127
di Neumann B.4
di operatori 215
trigonometriche di Fourier 132, 131
sistema:
di generatori 51
ortogonale 122
ortonormale 122
completo 124
soluzione debole 60, 158.3, 182.2
somma di funzioni definite
[quasi ovunque 42
sostegno:
di uno spazio topologico 4
di un sottospazio 9
sottodifferenziale 326
sottogradiente 326
sottoricoprimento 13

sottospazio:
metrico 112
normato 93
topologico 9
vettoriale 51, 87
generato da un sottoinsieme 51
normato 93
topologico 88
sottospazi supplementari 52, 120
spazi funzionali :
$\mathcal{B}(X \times X, Y)$ 211
$C^0(T)$ 56, 100, 114, 182 bis
$C^m(T)$ 56, 100, 114
$C^\infty(T)$ 56, 102
$C^{0,\lambda}(T)$ 56, 101, 114
$C^{m,\lambda}(T)$ 57, 101, 114
$C_0^0(\Omega)$ 57, 102
$C_0^m(\Omega)$ 57, 102
$C_0^\infty(\Omega)$ 58, 102
$\mathcal{C}^{m,p}(\Omega)$ 115, 164
$\mathcal{D}(\Omega)$ 102
$H^m(\Omega)$ 103, 114, 182
$H^{m,p}(\Omega)$ 103, 114, 182
$H_0^m(\Omega)$ 103, 114
$H_0^{m,p}(\Omega)$ 103, 114
l_p 55, 99, 114
l_∞ 55, 99
$\mathcal{L}(S, S_1)$ 203
$L^P(\Omega)$ 54, 70, 99, 114, 180, 188
$L^P(\Omega, \mu)$ 55, 100, 114
$L^p_{loc}(\Omega)$ 58, 102
$L^\infty(\Omega)$ 54, 99, 182
$\mathbf{R}^k$, $\mathbf{C}^k$ 70, 98, 114, 180
$W^m(\Omega)$ 61, 103, 114
$W^{m,p}(\Omega)$ 61, 103, 114
spazi vettoriali :
algebricamente isomorfi 62
normati:
isomorfi 140, 141
equivalenti 130, 140
spazio:
biduale 186
degli operatori
[lineari 69
[e continui 203
di Banach 113
riflessivo 187,189
uniformemente convesso 190
di Hausdorff 11
di Hilbert 113, 189
duale 69, 171, 173 (v. anche Duale)
euclideo 113
localmente compatto 15
localmente convesso 88
metrico 77, 95
normato 88, 95
completo 112
pre-hilbertiano 94, 95
prodotto 17, 53, 81, 93, 113
quoziente 93, 113
riflessivo 187
separabile 19
separato 11
strettamente convesso 323 nota ([1])
topologico 4, 95
vettoriale 49
di dimensione finita, infinita 50
normato 88
prodotto 53
quoziente 53
topologico 82, 95
localmente convesso 88
a dimensione finita 50,56,70,87,
145, 175,180,190,203bis([1])
spettrale (v. Rappresentazione)
spettro, di un operatore lineare 239
continuo 240
puntuale 240
residuo 240
successione :
convergente 19, 106, 110
debolmente 107, 197
puntualmente (oppure
[$_*$debolmente) 191
di Cauchy 112
di operatori:
convergente :
in norma 208
fortemente (oppure
[puntualmente
[in senso forte) 206
debolmente (oppure
[puntualmente
[in senso debole) 206
supplemento ortogonale 120
supporto 57,103 nota ([1])
compatto 57

T

teorema:
del grafico chiuso 142
della divergenza <11>, 4) pag.19
dell'alternativa 233, 235
della media 43
teorema di:
Ascoli-Arzelà 101
Banach 141
Banach-Caccioppoli 285
Beppo Levi 335
Bernstein [12]
Bolzano 23
Browder 286
Cantor 136
Eberlein-Smulian 200
Fubini 169 nota ([1])
Hahn-Banach 183
Hartags [12]
Heine-Cantor (V. Cantor)
Heine-Pincherle-Borel 14
Hilbert-Schmidt 246
Lax-Milgram 257
Nikodym 45, 149
Pitagora 96
di regolarità, per le
disequazioni variazionali 317
Riesz 177
Schauder 287
Sobolev 105, 304
Stampacchia 255
Stokes C.20
Tychonoff 287
Vitali-Lusin 35
Weierstrass 23
topologia 4
banale 4
compatibile con la struttura algebrica 82
debole, per uno spazio normato 195
$*$debole 194
della convergenza puntuale 30, 191
di base $\mathcal{B}$ 18
discreta 4
forte, per uno spazio normato 196
indebolita 194
indotta 9
più debole di un'altra 24
più fine di un'altra 24
più forte di un'altra 24
trasformata della δ di Dirac:
di Fourier 157
di Laplace 157

U

uguaglianza di Parseval 128

V

varietà lineare 52, 63
vero estremo inferiore 37
vero estremo superiore 36
vettore 50
colonna 64
vettori
ortogonali 96
linearmente indipendenti 50

Finito di stampare
nel mese di ottobre 2005
presso la Tipolitografia
Print Agency by I Farella Snc
NAPOLI