PRIMO INCONTRO CON LE EQUAZIONI INTEGRALI DI FREDHOLM

ESPLORAZIONE EURISTICA

II Edizione

Eric Ivar Fredholm, 1866-1927

(Per cultori di matematica ciclisti, con conoscenze a livello quinta Liceo Scientifico)

1. Introduzione

Tutti coloro che fanno le scuole medie incontrano le ***equazioni algebriche*** di *primo grado,* poi i sistemi di equazioni di primo grado, per cui vengono dati alcuni metodi di soluzione. Il più elegante è il metodo di **Cramer**, che ri-incontreremo, perché, anche se non è il metodo più efficiente, permette tuttavia euristicamente delle estensioni a concetti più avanzati. Questo metodo richiede l'impiego di determinanti: in questo saggio supporrò che metodo di Cramer e determinanti siano noti.

Poi arrivano le equazioni algebriche di *secondo grado*, con una loro formula risolutiva, che è facilmente dimostrabile usando il metodo del "completamento dei quadrati" o altro affine.

La soluzione è data da:

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

Sono numerosi coloro che bene o male la ricordano a memoria, assai più numerosi di quelli che la sanno ricavare. In quanto alle equazioni algebriche di grado superiore, è abbastanza noto che esistono formule risolutive per le equazioni di terzo e quarto grado ma credo che quelli che le ricordano o le sanno ricavare siano realmente pochini, anche tra i matematici. In quanto al fatto che non possono esistere formule basate sui radicali per risolvere equazioni di grado superiore al quarto (***Teorema di Ruffini-Abel***), penso che l'esistenza di questo teorema sia ignota ai più e incompresa dalla maggior parte di quelli che ne conoscono l'esistenza. Sia chiaro che tutte le soluzioni delle equazioni algebriche sono **numeri,** reali o immaginari.

Più avanti nel loro corso di studio, i nostri studenti, se frequenteranno il liceo scientifico, incontreranno le ***equazioni differenziali "alle derivate ordinarie"***, prima quelle del primo ordine e poi quelle del secondo ordine. Qui, in generale, **l'incognita da trovare è una funzione y di una variabile x,** in breve **y(x)**, e l'equazione è una relazione che lega la variabile indipendente, la funzione incognita e le sue derivate di vari ordini**:**

(1) $$F\left(x, y, y', y'' \dots y^{(n)}\right) = 0$$

L'equazione prende il nome dalla derivata di massimo ordine che vi compare. Ci sono metodi abbastanza numerosi, *anche se non esiste un metodo generale*, per risolvere le equazioni differenziali del *primo ordine,* cioè con derivate di ordine non superiore al primo, e metodi assai meno numerosi per risolvere certe classi di equazioni del secondo ordine. Infine, le equazioni di ordini superiori che possono essere risolte sono delle vere rarità: ma la natura sembra che si accontenti di sistemi che – almeno in prima approssimazione - richiedono solo equazioni lineari del secondo ordine. Il resto è terra riservata a pochi arditi esploratori. Chi vuol avere un'idea delle esplorazioni già fatte in questo vasto continente, può dare un'occhiata al classico testo di **E.L.Ince**, *Ordinary differential equations, 1926*). Qui vengono nominati, fino al 1926, meno di duecento di questi esploratori.

Alla fine del secondo corso universitario di Analisi Matematica (almeno quando studiavo io) arrivano le ***equazioni alle derivate parziali***, del primo e del secondo ordine. Il concetto non è differente dal precedente, ma la funzione incognita è una funzione di più di una variabile, per esempio x, y, v…. La (1) si trasforma così in:

(1b) $$F\left(x, y, v, \dots, \frac{\partial z}{\partial x}, \frac{\partial z}{\partial y}, \frac{\partial z}{\partial v} \dots \frac{\partial^n z}{\partial^m x\, \partial^l y\, \partial^k v \dots} \dots\right) = 0$$

(in cui $m+l+k\dots = n$)

Quando si parla di equazioni alle derivate parziali "della fisica matematica", ci si riferisce in genere ad equazioni al più del secondo ordine, sebbene ci sia qualche sparuto problema, come la vibrazione di piastre (non membrane infinitamente sottili), che richiede la soluzione di un'equazione del quarto ordine. E poi, per la massima parte del genere umano, c'è il deserto, esplorato anch'esso solo da matematici e fisici teorici.

A questo punto i curiosi forse si diranno: esistono equazioni differenziali, e va bene. Ma esistono anche ***equazioni integrali***? La risposta è, naturalmente, sì. In queste equazioni, la **funzione incognita** compare sotto il segno di integrale. Anzi, per complicare un poco le cose, esistono anche le equazioni cosiddette "***integrodifferenziali***", in cui compaiono le derivate della funzione incognita e in più essa compare sotto il segno di integrale. Queste equazioni sono utilizzate ad esempio nello studio dei circuiti elettrici (e affini, per esempio in neurologia), nonché in un soggetto di attuale interesse (aprile 2020), cioè l'epidemiologia, soprattutto ove si suddivida la popolazione in classi di età.

Tornando alle equazioni integrali, ne esistono di svariati tipi che portano i diversi nomi di chi le ha introdotte, o studiate o risolte per primo. Esse presentano il vantaggio "estetico" che, mentre le equazioni differenziali ci conducono a risolvere soddisfacentemente un problema purché in aggiunta all'equazione da risolvere specifichiamo le cosiddette condizioni al contorno, queste ultime possono essere incluse direttamente nelle equazioni integrali. Lo svantaggio è che le equazioni integrali sono in media assai più difficili da risolvere delle equazioni differenziali. Peccato! Ma non tutto è perduto: se le soluzioni "in forma chiusa" , cioè in termini di funzioni note, possono essere recondite, le soluzioni numeriche delle equazioni integrali sono spesso più accessibili di quelle delle equazioni differenziali.

Accertato che le equazioni integrali esistono e che presentano certi vantaggi (oltre che svantaggi) per la loro soluzione, resta la curiosità se il loro studio sia realmente utile, e non una pura "eleganza matematica". In effetti, le equazioni integrali (in particolare quelle del tipo di **Fredholm**) sono generalmente insegnate in Italia nel terzo corso di analisi, che non molti anni fa non era obbligatorio per gli studenti che non intendevano ottenere una laurea in matematica. Non era obbligatorio neppure per gli studenti del corso di fisica.

Un vero matematico, il problema dell'utilità di quello che studia o scopre, non se lo pone affatto. Ma talvolta si tratta solo di aver pazienza. Soggetti che sembrano essere gloriosamente inutili, d'improvviso possono acquistare un'importanza capitale in problemi fisici o ingegneristici nuovi. Un caso caratteristico è dato dal concetto di **autovalore**, che in verità presentava già un certo interesse fin da metà Ottocento, ma era destinato ad assumere un ruolo primario come strumento fondamentale della meccanica quantistica. Possiamo tranquillamente dire che senza equazioni agli autovalori la meccanica quantistica moderna non esisterebbe. Di qui, la teoria degli autovalori dilagò in vari campi della fisica teorica, tanto da permetterci di affermare che "*Il concetto di*

autovalore domina la fisica moderna". Ora, le equazioni integrali sono uno dei mezzi di elezione per il calcolo degli autovalori di un sistema che ne possegga.

Tornando alle equazioni integrali, posso ora enumerare alcuni campi di interesse delle equazioni integrali:

i) Le equazioni integrali hanno anzitutto un *valore storico*, in quanto sono all'origine della cosiddetta "analisi funzionale", ai primi dell'Ottocento.

ii) Esse hanno in secondo luogo un *valore teorico*, affiancando le equazioni differenziali. Ad esempio, l'esistenza e unicità delle soluzioni delle equazioni differenziali (date opportune condizioni al contorno), sono normalmente dimostrate utilizzando un'equazione integrale (Teorema attribuito a **Émile Picard, Ernst Lindelöf, Rudolf Lipschitz and Augustin-Louis Cauchy** – il quale , come il solito, fu essenzialmente il primo a dimostrarne l'assunto). Il teorema è normalmente dimostrato per un'equazione differenziale ordinaria del primo ordine, e può essere esteso senza eccessiva difficoltà a un sistema di n equazioni del primo ordine. Ma, come è ben noto a chi è noto, *un sistema di n equazioni differenziali del primo ordine può essere ridotto a un'unica equazione differenziale di ordine n* (e viceversa).

iii) Esse forniscono il *mezzo naturale per trattare diversi problemi della fisica*. La formulazione integrale delle equazioni di Maxwell dell'elettromagnetismo è forse l'applicazione più nota, ma non mancano applicazioni ai problemi di trasferimento radiativo, acustica etc.. Ci sono inoltre problemi che non sono trattati agevolmente per mezzo di equazioni differenziali, ma sono accessibili per mezzo delle equazioni integrali, in quanto il comportamento della funzione incognita in un punto P, dipende dai valori della funzione a una distanza finita da P, e non solo a distanza infinitesima.

(iv) Come vedremo, talune equazioni integrali sono il *modo naturale per trattare gli autovalori di determinati sistemi*, che corrispondono alle frequenze proprie di sistemi oscillanti, carichi critici di sistemi rotanti e altri problemi più complessi, ponendosi almeno idealmente, come si è detto, al centro della fisica teorica moderna.

2. Esempio di problema meccanico che genera un'equazione integrale.

Si consideri un'asta rotante, come ad esempio l'albero dell'elica di una nave. Se l'albero viene posto in rotazione partendo da velocità angolare zero, si noterà che, raggiunta una certa velocità di rotazione critica, in genere assai inferiore alla velocità massima consentita, esso sarà soggetto a pericolose vibrazioni grosso modo normali all'asse di rotazione. Aumentando la velocità di rotazione, però, queste oscillazioni si attenueranno fino a scomparire, almeno fino a che non sia raggiunta una seconda velocità critica, e così via. Oggi (che si viaggia assai meno in nave) il fenomeno è meno percepibile, tanto più che vi sono accorgimenti per ridurlo.

Per vedere come si possa giungere a determinare queste velocità critiche per mezzo del calcolo, possiamo considerare, anzitutto, un problema di statica, cioè il comportamento di una trave (l'albero dell'elica), di lunghezza 1, appoggiata o fissata agli estremi. "Intuitivamente" accetteremo il fatto che esiste una *"funzione d'influenza"*, che indicheremo con G(x,y), la quale dà lo spostamento verticale z all'ascissa x, se un carico unitario verticale, cioè lungo z, è applicato all'ascissa y - comunque la trave sia fissata o appoggiata agli estremi.

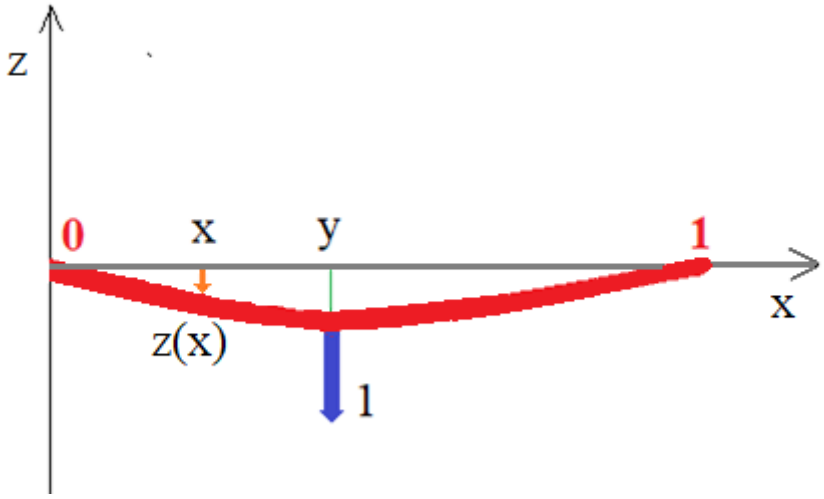


Fig.1

Ciò che interessa è che questa funzione rappresenta la risposta (statica) di una trave elastica ad una distribuzione di forze. Se per esempio avessimo un sistema costituito da una massa pesante e una molla, avremmo che in quel punto, la forza è proporzionale all'allungamento. Come chiaramente annunciò Hooke: *ceiiinossstuv*, anagramma di « Ut tensio, sic vis", ovvero $\boldsymbol{F} = -k\,\boldsymbol{\Delta l}$, dove k è la costante elastica della forza.

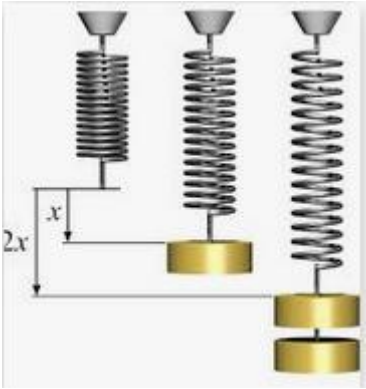


Fig.2

Svjo / CC BY-SA (https://creativecommons.org/licenses/by-sa/3.0)

Dove k è la costante elastica della forza. Nel caso in cui la forza sia il peso, potremmo

$$\Delta l = \left(-\frac{1}{k}\right) F$$

In embrione, il termine (-1/k) è il prototipo di una *funzione di influenza*, ridotta ai minimi termini, in quanto trave, carico e "flessione" sono ridotti ad un unico punto, cioè x=y. A questo punto possiamo "intuire" che, se invece abbiamo una trave elastica continua, l'effetto di una forza, o carico, che agisce in un unico punto y è sentito in modo diverso nei vari punti x della trave.

Ancora più complicato è il caso in cui non agisce una sola forza in un solo punto, ma una distribuzione di forze. Ma qui ci soccorre un altro principio semi-intuitivo, quello della "sovrapposizione degli effetti", che ci dice che se l'asta è sottoposta a un sistema di forze f_i in direzione z, applicate in $\boldsymbol{n}$ punti $\boldsymbol{i}$ dell'asta, lo spostamento z in x è dato dalla somma di tutti gli effetti dovuti tutte le varie forze applicate nei vari punti $\boldsymbol{i}$ della trave. Naturalmente, ciò significa che se il sistema di forze è applicato *con continuità* sull'asta potremo passare "euristicamente" al limite per $n = \rightarrow \infty$, e magari continuo, con un accorgimento che useremo spesso più avanti:

(2) $$z(x) = \sum_{i=1}^{n} G(x,i) f_i \, \Delta i \;\rightarrow \int_0^1 G(x,y) f(y) dy$$

in cui la sommatoria diventa un integrale, la variabile discreta i della sommatoria diventa y, e il termine Δi, che vale 1 nella sommatoria, diventa l'infinitesimo dy nell'integrale. Questo trucco, che un matematico considererebbe abbietto, sarà alla base del saggio che segue. Matematico avvisato è mezzo salvato.

Per la cronaca, la funzione G(x, y) andrebbe determinata semiempiricamente e semiteoricamente caso per caso. **Il teorema di reciprocità di Maxwell** (1864, poi generalizzato da Betti) comunque ci aiuta assicurandoci che G(x,y)= G(y,x), cioè lo spostamento lungo z in x dovuto a un carico unitario applicato in y è eguale allo spostamento lungo z in y , dovuto a un carico unitario applicato in x.

Noi abbiamo sfruttato il concetto, abbastanza intuitivo, che una trave a riposo non può essere infinitamente rigida. Essa è più o meno dotata di elasticità. Ma un effetto interessante di questo fatto è che se la trave è allontanata dalla posizione di riposo (z(x) =0 identicamente), una forza elastica di richiamo tende a riportarla alla posizione di riposo, ciò che avviene dopo oscillazioni che sono presto smorzate se non corrispondono a certe determinate frequenze critiche, che danno luogo a "oscillazioni libere", cioè non sollecitate da forze, come, ad esempio, le vibrazioni della corda di una chitarra, che può vibrare una volta pizzicata, senza che altre forze intervengano a sostenere la vibrazione.

Se ora poniamo l'asta in rotazione con velocità angolare ω intorno all'asse x, sappiamo che esiste una forza centrifuga, nulla sull'asse di rotazione, la quale può esercitare un carico traversale sui baricentri delle sezioni dell'asta, se essi sono per qualche motivo allontanati a una piccola distanza dall'asse, o se lo sono per difetto di costruzione. In tal caso, il carico che agisce sul tronco di asta fra le ascisse x e x+dx ha l'espressione:

(3) $$p(x)dx = \omega^2 z(x)dm = \omega^2 z(x)\ \mu(x)dx$$

in cui z(x) è l'allontanamento del baricentro dell'asta dalla posizione di riposo, e dm è dato dal prodotto della densità di massa $\mu(x)$ nel punto x, per la lunghezza dx.

Se questo carico dovuto alla forza centrifuga equilibra la forza elastica di richiamo, la trave non può ritornare permanentemente alla posizione di riposo, ma oscillerà con la sua frequenza propria. Avremo allora delle oscillazioni libere trasversali al moto di rotazione, se la frequenza di rotazione eguaglierà la frequenza di una delle oscillazioni libere proprie dell'asta.

E' quindi "intuitivo" che si creino situazioni di equilibrio instabile quando la rotazione intorno all'asse x provoca sul segmento dx considerato un carico esattamente eguale a quello che in condizioni statiche provoca lo stesso spostamento e mantiene le stesse oscillazioni libere trasversali, cioè quando

$$f(x)dx = \omega^2 z(x)\ \mu(x)dx$$

Ovvero:

(4) $$f(x) = \omega^2 z(x)\ \mu(x)$$

Cioè quando "lo spostamento del baricentro, dovuto alla rotazione, genera una forza che in condizioni statiche è in grado di generare il medesimo spostamento". (Credo che di qui venga la parte "auto-" del termine autovalore, almeno in italiano. In tedesco e inglese l'autovalore sarebbe piuttosto un "valore proprio (eigenvalue)".)

In tal caso, l'equazione integrale omogenea:

(5) $$z(x) - \omega^2 \int_0^1 G(x,y)\mu(y)\ z(y)dy = 0$$

può ammettere soluzioni "proprie", cioè diverse dalla soluzione banale z (x) = 0 (nessuna oscillazione). Tali soluzioni, essendo lungo l'asse z, sono oscillazioni trasversali all'asse di rotazione.

Il problema meccanico, di identificare le velocità di rotazione critiche, è risolto trovando il valore di ω^2 tale che l'equazione omogenea (5) abbia soluzioni proprie. Tale equazione è detta **equazione integrale lineare omogenea di Fredholm del secondo tipo.**

Ma è vero che tutto ciò è così intuitivo? Ci sono troppi concetti "intuitivi" in questa introduzione! Per me l'insieme è assai poco intuitivo. Tutte le mie equazioni integrali hanno sempre avuto un fondamentale sassolino nella scarpa: non mi era chiaro perché la situazione dovesse portare, ad esempio, alla frattura dell'asse dell'elica, ciò che, soprattutto in navi di media dimensione, poteva portare al rapido affondamento della nave. Si parla di "entrare in risonanza". Ma che vuol dire esattamente? L'idea di risonanza implica che in qualche modo esistano almeno due frequenze identiche. Certo, in qualche modo ci si arriva e qualcosa succede, ma cosa? E come?

Formalmente, tutto appare molto semplice. Ma ho dovuto studiare a lungo come rendermi intuitivo il concetto, e non ho trovato di meglio dell'esempio seguente. Se lo si comprende, e si comprende come possa seguirne un processo disastroso, poi, secondo me, si può rileggere con maggiore sicurezza il testo che ho già scritto, a partire da quando parlo di mettere in rotazione la trave, qualche riga prima dell'equazione (3).

Sia dunque il modello elementare di un volano di massa m montato su un asse elastico di massa trascurabile rispetto a quella del volano. Il volano, per qualche motivo, è lievemente eccentrico, con eccentricità e (distanza del baricentro G dal centro O – a riposo). Vogliamo trovare la velocità critica di rotazione dell'asse se la frequenza *naturale* del volano è $\omega_n = \sqrt{(k/m)}$, in cui k è la costante di elasticità dell'asse considerato come trave elastica. Per frequenza naturale intendiamo la frequenza del movimento oscillatorio del volano se allontaniamo il punto R dall'asse verticale. Sia R il centro di rotazione, G il baricentro del volano, O il centro geometrico del rotore. Per eccentricità si intende la distanza $e = (OG)$ (vedi figura 3b). Il punto G è per noi nascosto nel volano; per il punto R passa l'asse elastico; O, centro geometrico del volano *a riposo*, è un punto fisso che deriva da quando il sistema fu costruito. A riposo, R e O coincidono. In un'oscillazione ***trasversale*** il punto R visibilmente si allontana dalla verticale, il punto G, che già non è sulla verticale, oscilla ma resta nascosto nel volano, il punto O resta sulla verticale, ma non è più fisso nel volano.

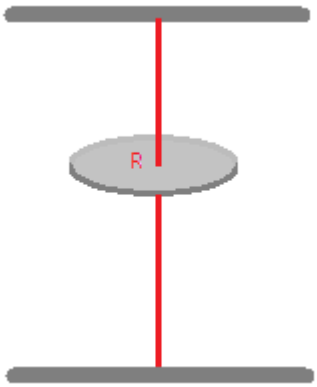


Fig. 3 a

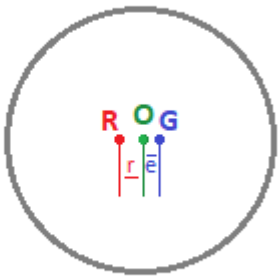


Fig. 3 b

In Fig.3 a la trave rotante è stata schematizzata in due porzioni: un volano di massa m, infinitamente sottile; e l'asse elastico (in rosso) intorno a cui esso ruota, con massa zero. Esso è pure infinitamente sottile.

Fig. 3 b riporta una visione zenitale del volano, con le distanze r ed e molto ingrandite.

Ad ogni istante la forza di richiamo dovuta all'elasticità dell'asse è F = kr , dove r è la distanza (RO). Invece, la forza centrifuga dovuta al fatto che la ruota non è equilibrata, è m (e+r) ω^2. Queste due forze devono farsi equilibrio, cioè:

(6) $$kr = m\,(e+r)\omega^2 \;\rightarrow \left(\frac{k}{m}\right) = \frac{(e+r)\omega^2}{r}$$

$$\omega_n^2 r - \omega^2 r = e\omega^2 \rightarrow r\,(1 - \frac{\omega^2}{\omega_n^2}) = e\frac{\omega^2}{\omega_n^2}$$

Da cui,

(7) $$r = \omega^2 \frac{\frac{e}{\omega_n^2}}{(1-\frac{\omega^2}{\omega_n^2})}$$

Ne segue che r tende a diventare infinito quando ω si avvicina a ω_n . Ma, nel nostro esempio, r = infinito significherebbe che il sistema è volato in pezzi. Una barra elastica omogenea non si comporterebbe diversamente: qui è stata riassunta in un volano con asse (di massa trascurabile) elastico. L'unico fattore (del resto sempre presente) che potrebbe evitare questo disastro è lo smorzamento dovuto a attrito o altro. In tal caso ci sarebbe un addendo in più al denominatore, che non permetterebbe a questo di annullarsi, e la velocità critica potrebbe essere oltrepassata senza danni (ma la frattura può avvenire anche in presenza di smorzamento, se questo non è sufficiente). Qui è dunque ovvio qual è la velocità di rotazione critica: è la velocità ω eguale alla frequenza "naturale" di oscillazione del volano. In una trave la situazione non è diversa, ma è più complicata, e la frequenza di oscillazione naturale della trave è uguale all'autovalore dell'equazione integrale (5). Rileggendo in questa luce la (5) vediamo che se per un certo valore di ω la z(x) va all'infinito, ciò vuol dire che la trave si è spezzata.

Ma torniamo alle nostre equazioni integrali. Più in generale, l'e quazione di Fredholm del secondo tipo include il caso non omogeneo della forma:

(8) $$\varphi(x) - \lambda \int_0^1 K(x,y)\varphi(y)dy = f(x)$$

Qui, per adeguarci alle forme comunemente usate, si è posto

$$z(x) \rightarrow \varphi(x);\;\; \omega^2 \rightarrow \lambda\,(>\mathbf{0});\; G(x,y)\mu(y) \;\rightarrow\; K(x,y)$$

Detto per inciso, possiamo trasformare l'equazione omogenea (5) in modo redditizio moltiplicando ambo i membri per $\sqrt{\mu(x)}$ ed introducendo la funzione

$$\psi(x) = \sqrt{\mu(x)}\, z(x)$$

la quale, sotto integrale , diviene

$$\psi(y) = \sqrt{\mu(y)}\, z(y)$$

Avremo così

$$\psi(x) - \omega^2 \int_0^1 G(x,y)\sqrt{\mu(x)}\,\sqrt{\mu(y)}\,\psi(y)dy = 0$$

il cui nucleo $K(x,y) = G(x,y)\sqrt{\mu(x)}\,\sqrt{\mu(y)}$ è evidentemente simmetrico rispetto allo scambio di x e y, *il che garantisce in generale infiniti autovalori*, ciò che non è sempre vero di un nucleo asimmetrico, che può anche non avere autovalori. Ma di questo non ci occuperemo, se non assai di striscio.

Per arrivare da un punto di vista matematico al concetto di equazione integrale, io seguirò ora, come annunciato fin dal titolo, un metodo euristico, che dovrebbe permettere di imparare qualcosa, in modo spero indolore, anche su altri concetti.

Per questo primo incontro, mi limiterò alla situazione più semplice, con soli numeri reali.

Ma prima occorre avere qualche nozione su matrici, determinanti e questioni connesse. Speravo che si potesse fare a meno di questa parte, o lasciarla in appendice, ma devo dire che ho notato che le conoscenze di questi concetti asono assai poco diffuse. Senza scrivere un trattato a pieno titolo, la mia ambizione sarebbe quella di rendere intuitivi i concetti principali.

COMPLEMENTO A.

NOZIONI EL:EMENTARI SU MATRICI, DETERMINANTI, REGOLA DI CRAMER, AUTOVALORI, AUTOVETTORI/AUTOFUNZIONI.

Che cosa sono le matrici?

Semplicemente delle tabelle, che per noi avranno sempre un ugual numero di righe e di colonne con due sola eccezioni (anche se si potrebbe essere più generali). Le chiameremo matrici 2 x 2 , 3 x 3, 4 x 4 eccetera. Una matrice 1 x 1, ha una riga ed una colonna, e quindi è un numero isolato. Quello che avrà di buono il prodotto di matrici (che risrva almeno una sorpresa) è che include anche il caso del prodotto di due numeri ordinari.

L'eccezione che avevo menzionato è che considereremo anche matrici costituite da una sola colonna di elementi:

$$\begin{pmatrix} a \\ b \\ c \end{pmatrix}$$

Come discuteremo meglio in seguito, una matrice colonna la chiameremo **vettore**.
I vari oggetti che compaiono in ogni casella della tabella (o della colonna) sono gli "elementi" di matrice.

Per quanto riguarda le quattro operazioni, bisogna dire che sono tutte possibili, anche se sotto qualche condizione.
Addizione e sottrazione non sono niente di speciale: si esegue l'addizione elemento per elemento.

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} + \begin{pmatrix} A & B \\ C & D \end{pmatrix} = \begin{pmatrix} a+A & b+B \\ c+C & d+D \end{pmatrix}$$

e la sottrazione allo stesso modo.
Le cose si complicano quando si parla di prodotto: la regola è a prima vista complicata ed è facile confondersi.
Regola del prodotto di matrici, per esempio 2 x2:

$$\begin{pmatrix} \boldsymbol{a} & \boldsymbol{b} \\ \boldsymbol{c} & \boldsymbol{d} \end{pmatrix} \cdot \begin{pmatrix} \boldsymbol{A} & \boldsymbol{B} \\ \boldsymbol{C} & \boldsymbol{D} \end{pmatrix} = \begin{pmatrix} \boldsymbol{x} & \boldsymbol{y} \\ \boldsymbol{w} & \boldsymbol{z} \end{pmatrix}$$

La regola, evidentemente, deve permetterci di calcolare x,y,w,z.

Ricetta:
Affettare la prima matrice in due fette orizzontali e la seconda in due fette verticali.

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \cdot \left(\begin{array}{c|c} A & B \\ C & D \end{array}\right)$$

Quindi ottenere il risultato eseguendo le seguenti operazioni "riga per colonna":

- x = a A + b C: cioè primo elemento della prima riga della colonna di sinistra per il primo elemento della prima colonna della matrice di destra **a cui si somma** il prodotto del secondo elemento della prima riga della matrice di sinistra per il secondo elemento della prima colonna della matrice di destra.

- y = a B + b D
- w = c A + d C
- z = c B + d D

Però la regola è più complicata a descriversi che a mettersi in pratica, anche se non sembra esserci alcuna ragione perché la regola debba essere così.
Ma perché una regola così complicata? Semplicemente perché è utile per certe

applicazioni.

Intanto esiste una spiegazione immediata:
Ad esempio, si supponga che , essendo noti a,b,c,d, eseguiamo la cosiddetta"sostituzione lineare"

x= au + bv, y = cu + dv.

Poi supponiamo che a loro volta

u=AU+BV, v=CU+DV.

Ora vogliamo scrivere x e y in termini di U e V, saltando le u e v.
Allora x = a(AU+BV)+b(CU+DV) =raccogliendo le U e le V = **(aA+bC)** U **+ (aB+bD)** V
Come esercizio si può trovare anche y in termini di U e V.
Il risultato è che ritroviamo ritroviamo i nostri quattro risultati del prodotto righe per colonna delle matrici

$$\begin{pmatrix} \boldsymbol{x} \\ \boldsymbol{y} \end{pmatrix} = \begin{pmatrix} \boldsymbol{a} & \boldsymbol{b} \\ \boldsymbol{c} & \boldsymbol{d} \end{pmatrix} \cdot \begin{pmatrix} \boldsymbol{A} & \boldsymbol{B} \\ \boldsymbol{C} & \boldsymbol{D} \end{pmatrix} \cdot \begin{pmatrix} \boldsymbol{U} \\ \boldsymbol{V} \end{pmatrix}$$

Applicando la regola complicata che ho appena dato, al prodotto di due matrici di una riga ed una colonna, per esempio

(a) x (b) troviamo ancora (ab).

Infatti il primo - ed unico - elemento della prima riga della prima colonna del risultato è dato ddal prodotto del primo e unico elemento della prima riga della matrice di sinistra per il primo e unico elemento della prima colonna della matrice di destra.
In altre parole, la nuova moltiplicazione include ed estende, ma non contraddice, il vecchio concetto di moltiplicazione.

Una semplice applicazione è come scrivere un sistema lineare in termini di matrici e vettori.
Sia il sistema

$$\begin{cases} Ax + B\,y = a \\ Cx + Dy = b \end{cases}$$

Sia la matrice

$$M = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$$

Applicando la regola del prodotto righe per colonne, potremmo scrivere:

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} a \\ b \end{pmatrix}$$

Oppure, in modo abbreviato:

$$M\,\mathbf{u} = \mathbf{v}$$

Si può introdurre un nuovo formalismo. Introducendo nuovi formalismi occorre solo abituarsi. Invece di chiamare una matrice

$$M = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$$

Possiamo introdurre indici:

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix};\ \mathrm{B} = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix}; M = \begin{pmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{pmatrix}$$

Con la convenzione che il primo indice indica la riga, il secondo la colonna.
Il prodotto di due matrici sarebbe quindi scritto:

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}\begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} = \begin{pmatrix} a_{11}b_{11} + a_{12}b_{21} & a_{11}b_{12} + a_{12}b_{22} \\ a_{21}b_{11} + a_{22}b_{21} & a_{21}b_{12} + a_{22}b_{22} \end{pmatrix} = \begin{pmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{pmatrix}$$

Notiamo che in questo caso (matrice 2 x 2) ogni elemento della matrice prodotto, M, è dato da due addendi composti da prodotti in cui il primo e l'ultimo indice (in rosso) sono gli indici previsti per l'elemento di M, mentre gli indici interni (in verde) sono eguali e sono 1,1 nel primo addendo, e 2,2 nel secondo.

Introducendo il simbolo di sommatoria, e pensandoci un momento, abbiamo che

$$M_{ij} = \sum_{k=1}^{2} a_{ik}\, b_{kj}$$

Se per divertimento si esegue una moltiplicazione di due matrici quadrate 3 x 3, si vede che nulla cambia nella formula appena data, eccetto il fatto che la somma va da k=1 a k=3, e allo stesso modo gli indici i e j vanno da 1 a 3, producendo una matrice 3 x 3 in cui ogni elemento è dato da tre addendi.
Per esempio, si può scrivere subito:

$$M_{32} = a_{31}b_{12} + a_{32}b_{22} + a_{33}b_{32}$$

In generale, quindi, per una matrice n x n :

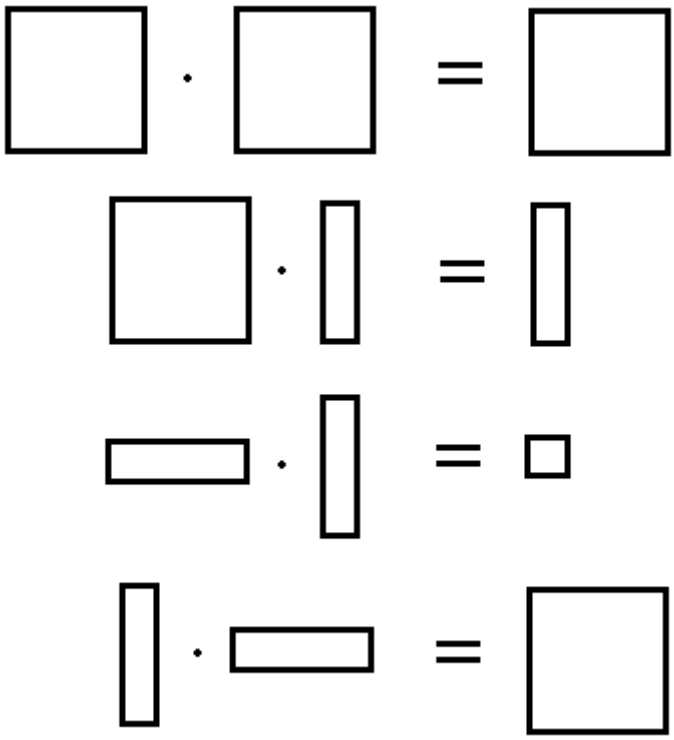
$$(AB)_{ij} = \sum_{k=1}^{n} a_{ik}\, b_{kj}$$

Ancora un passo, escogitato dai fisici e (almeno da principio) non visto tanto bene dai matematici: nulla ci vieta di scrivere:

$$(AB)_{ij} = \sum_{k=1}^{n} a_{ik}\, b_{kj} \rightarrow a_{ik}\, b_{kj}$$

evitando di scrivere la lunga sommatoria, ma assumendo che essa si svolge da 1 a n (e il fisico deve sapere a priori quanto vale n) ***sommando sugli indici ripetuti***. E? ovvio che gli indici ripetuti, *dato che completando la somma essi scompaiono*, possono esse j,j oppure n,n o qualsiasi altra lettera dell'alfabeto.

Enrico Fermi presentava nelle sue lezioni un riassunto grafico dei prodotti che coinvolgevano matrici quadrate e vettori:



(Il diagramma più importante in questo contesto è il secondo, il meno immediato da ricostruire è l'ultimo, ma non lo utilizzeremo)

Fig.4

La regola del prodotto diventa così abbastanza automatica, e si può immaginare come applicarla al prodotto di due matrici 3x3. Ma il prodotto di matrici riserva qualche sorpresa. Noi ci accontenteremo di matrici semplici. Per esempio:

$$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$$

Mentre, moltiplicando in ordine inverso

$$\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

Qui insomma era tutto il succo di questo esercizio. ***I due risultati sono diversi.***
A x B non è uguale a B x A.

Ci sono tre classi di studenti di matematica. Quelli che dicono: "E a me che me ne...?" e si dedicano ad altro pensando di aver perso il proprio tempo (sono di norma la maggioranza). Poi ci sono quelli che dicono: "Ah che bello! Adesso mi metto a giocare con le matrici". Poi si vanno a cercare dei libri o delle pagine su Internet sul soggetto e divengono dei maghi. E infine ci sono quelli che dicono: ma tutto ciò serve a qualcosa? Vedremo.

(Per rendere la presentazione minimale resteremo nel campo reale e tratteremo solo matrici quadrate e vettori, cioè matrici a una colonna/riga, preferibilmente in due, massimo 3 dimensioni)

Matrici inverse, regola di Cramer.

Si prenda come esempio una matrice quadrata del tipo più semplice possibile.
Sia la matrice

$(A1)$ $$M = \begin{pmatrix} 1 & 4 \\ 1 & 1 \end{pmatrix}$$

che risulterebbe, per esempio, dal sistema di equazioni lineari:

$$\begin{cases} x + 4\,y = a \\ x + y = b \end{cases}$$

O anche:

$$\begin{pmatrix} 1 & 4 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} a \\ b \end{pmatrix}$$

Che si può scrivere in modo abbreviato:

$$\mathrm{M}\,\mathbf{u} = \mathbf{v}$$

Accanto ai vari metodi di soluzione, il più elegante e quello che permette estensioni nei campi più avanzati, è il **metodo di Cramer** (dal ginevrino Gabriel Cramer, 1750), che, pur cucinato in varie salse, si basa sullo spunto formale che

$$\boldsymbol{u} = \frac{1}{M}\,\boldsymbol{v} \;=\; M^{-1}\boldsymbol{v}$$

Dove M^{-1} è la "matrice inversa".

Divisione di matrici, matrice inversa.

Per ottenere la regola di Cramer in modo utilizzabile, occorre però conoscere la regola della divisione di matrici. E questo richiede, per chi non la conosce, una digressione, la più lunga di questo saggio.

Bisogna dire che la divisione è sempre la più difficile delle quattro operazioni: non la potenza, che può essere svolta come una successione di prodotti, in cui l'ordine è irrilevante:

$$A^3 = A^2\ A = A\ A\ A = A\ A^2$$

E quindi, dirà qualcuno, anche uno sviluppo in serie di matrici non dovrebbe presentare troppe difficoltà. Lascio chi ha uno spirito inquisitivo il piacere di continuare su quella strada, che non è breve né priva di interesse (come esprimerebbe exp(A)? gli interesserebbe provare con una matrice 2 x2? E sin (A)?)

Ma intanto diciamo che per noi l'essenziale è saper costruire la matrice inversa, A^{-1}, tale che $A\ A^{-1} = I$, la matrice identità:

$$I = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Per gli elementi della matrice identità sovente si usa il simbolo, detto di Kronecker, tale che $\delta_{ij,} = 1$, se $i = j$ e $\delta_{ij,} = 0$ se $i \neq j$, simbolo che non dipende dal numero n nella matrice n x n.

La divisione $B/A = B\ A^{-1}$, ma potrebbe essere anche $A^{-1}B$, sperando che l'ordine sia irrilevante. ***Non lo è necessariamente.*** Se abbiamo l'equazione $B= XA$, la soluzione la si ottiene moltiplicando a destra per A^{-1}, ottenendo $B\ A^{-1}= X\ AA^{-1} = XI = X$; mentre se l'equazione è $B = AX$, la soluzione la si ottiene moltiplicando a sinistra per A^{-1}, ottenendo $A^{-1}B = A^{-1}\ AX = IX = X$. Poiché il prodotto non commuta sempre, le due X così ottenute possono essere differenti.

Tutto sta quindi nel saper costruire la matrice inversa, tale che $M\ M^{-1} = I = M^{-1}M$, dove I è la matrice identità.
Vediamo il caso di una matrice 3 x 3.
Vogliamo cioè l'inverso della matrice

(A2) $$M = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

Per arrivare a una formulazione di M^{-1} coerente con quanto già scritto, il metodo che sceglieremo è quello di passare attraverso il concetto di determinante. Come è noto, a ogni matrice **n x n**, che è una tabella di numeri, si può associare il suo determinante, che è un numero singolo. Questo in generale è dato dalla somma di tutti i prodotti possibili di n elementi presi da **diverse righe e diverse colonne**, ciascuno con un segno opportuno. Se cerchiamo questi elementi direttamente con ordine, mettendo le righe nell'ordine naturale (1,2,3…) e le colonne come possono, abbiamo i sei prodotti

$$a_{11}a_{22}a_{33},\ a_{11}a_{23}a_{32},\ a_{12}a_{21}a_{33},\ a_{12}a_{23}a_{31},\ a_{13}a_{21}a_{32}, a_{13}a_{22}a_{31}$$

I nove elementi della matrice compaiono ciascuno due volte, ma in gruppi di tre, perciò il numero di addendi nel determinante è (9 x 2)/3 = 6.
I segni vanno dati secondo che gli indici delle colonne (i secondi indici) appartengono a una classe pari (segno +) o dispari (segno -) di permutazioni, rispetto a quella naturale data dagli indici di riga (i primi indici). La classe è pari o dispari se i numeri si discostano dall'ordine naturale (1,2,3) un numero pari o dispari di volte. Ad esempio , il primo prodotto è di classe pari, perché anche gli indici delle colonne si seguono nell'ordine 1,2,3 - segno più; nel secondo prodotto gli indici delle colonne si seguono nell'ordine 1,3,2, con una inversione - segno meno; l'ultimo ha tre inversioni 3,2,1 (3 prima di 2, 3prima di 1, 2 prima di 1)- segno meno. Il determinante D, trovato in quello che chiameremo "**il metodo diretto**", quindi vale:

(A3) $$D = \boldsymbol{a_{11}a_{22}a_{33}} - \boldsymbol{a_{11}a_{23}a_{32}} - \boldsymbol{a_{12}a_{21}a_{33}} + \boldsymbol{a_{12}a_{23}a_{31}} + \boldsymbol{a_{13}a_{21}a_{32}} - \boldsymbol{a_{13}a_{22}a_{31}} =$$
$$a_{11}(a_{22}a_{33} - a_{23}a_{32}) - a_{12}(a_{21}a_{33} - a_{23}a_{31}) + a_{13}\,(a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31})$$

Si ossrvi la seconda espressione: i coefficienti degli elementi della prima riga della matrice, in rosso, sono dei determinanti 2x2, in rosso, detti "complementi algebrici" degli elementi a_{11} , a_{12} e a_{13}, che si ottengono dal determinante 3 x 3 cancellando la riga e la colonna che si incrociano negli elementi della prima riga a_{11} , a_{12} e a_{13} rispettivamente, come nella seguente figura:

$$\begin{vmatrix} \boxed{a_{11}} & a_{12} & a_{13} \\ a_{21} & \boxed{a_{22}} & \boxed{a_{23}} \\ a_{31} & \boxed{a_{32}} & \boxed{a_{33}} \end{vmatrix} + \begin{vmatrix} a_{11} & \boxed{a_{12}} & a_{13} \\ \boxed{a_{21}} & a_{22} & \boxed{a_{23}} \\ \boxed{a_{31}} & a_{32} & \boxed{a_{33}} \end{vmatrix} + \begin{vmatrix} a_{11} & a_{12} & \boxed{a_{13}} \\ \boxed{a_{21}} & \boxed{a_{22}} & a_{23} \\ \boxed{a_{31}} & \boxed{a_{32}} & a_{33} \end{vmatrix}$$

In quanto al segno di ogni gruppo di termini, esso è dato dalla somma degli indici degli elementi a cui si riferisce ogni minore.

Se la somma è pari, il segno è positivo, se è dispari , il segno è negativo. Ne risulta quindi la tabella:

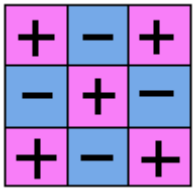


Fig.5

Ma possiamo anche raggruppare e riordinare i vari addendi in modo diverso, come ad esempio:

$a_{11}(a_{22}a_{33} - a_{23}a_{32}) - a_{21}\,(a_{12}a_{33} - a_{13}a_{32}) + a_{31}(a_{12}a_{23} - a_{13}a_{22})$

In questo caso, gli stessi sei addendi sono ordinati secondo i tre elementi della prima colonna. Di nuovo abbiano tre termini dati dal prodotto di un elemento della prima colonna per il "suo complemento algebrico" ottenuto cancellando la riga e la colonna che si incrociano nell'elemento di partenza, ciascuno ol segno dato dalla medesima tabella.

$$\begin{vmatrix} \boxed{a_{11}} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} + \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ \boxed{a_{21}} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} + \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ \boxed{a_{31}} & a_{32} & a_{33} \end{vmatrix}$$

Allo stesso modo il determinante può essere sviluppato secondo gli elementi di una stessa riga qualsiasi o di una stessa colonna qualsiasi.

Si può notare che se il determinante fosse 4 x4, si potrebbe calcolarlo (oltre che col metodo diretto) sviluppandolo secondo gli elementi di una stessa riga o di una stessa colonna, Ciascuno elemento risulterebbe moltiplicato per il suo complemento algebrico, che questa volta sarebbe un determinante 3 x3, il cui segno sarebbe determinato da un diagramma analogo alla figura 5. Poi, ogni complemento algebrico andrebbe a sua volta sviluppato come se fosse un qualsiasi determinante 3 x 3, per un totale di 12 determinanti 3 x3, ciascuno corrispondente a tre determinanti 2 x2. Totale, 36 determinanti 2 x 2: un bel lavoro, che costituisce la cosiddetta **"Regola di Laplace",** che qui non è stata dimostrata, ma solo mostrata con un paio di esempi. Va notato che per lo sviluppo di determinanti maggiori esistono regole pratiche assai più efficienti, che qui non ci riguardano.

Consideriamo ora due matrici:

$$a = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \quad e \quad b = \begin{pmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{pmatrix}$$

Abbiamo, applicando il "metodo diretto", che il determinante della seconda matrice è data da:

$$\mathbf{det}(\boldsymbol{b}) = \boldsymbol{b_{11}b_{22}b_{33}} - \boldsymbol{b_{11}b_{23}b_{32}} - \boldsymbol{b_{12}b_{21}b_{33}} + \boldsymbol{b_{12}b_{23}b_{31}} + \boldsymbol{b_{13}b_{21}b_{32}} - \boldsymbol{b_{13}b_{22}b_{31}}$$

Supponiamo che la matrice b sia la matrice a con le prime due colonne scambiate, cioè una matrice in cui: b(1,1)= a(1,2); b(1,2) = a(1,1) ; b(1,3)= a(1,3); b(2,1) = a(2,2); b(2,2) = a(2,1); b(2,3) = a(2,3); b(3,1) =a(3,2); b(3,2) = a(3,1); b(3,3) = a(3,3).

Avremmo allora:
$\boldsymbol{det}\ (\boldsymbol{b}) = \boldsymbol{a_{12}a_{21}a_{33}} - \boldsymbol{a_{12}a_{23}a_{31}} - \boldsymbol{a_{11}a_{22}a_{33}} + \boldsymbol{a_{11}a_{23}a_{32}} + \boldsymbol{a_{13}a_{22}a_{31}} - \boldsymbol{a_{13}a_{21}a_{32}}$
Paragonando a
$D\ (= \det a) = \boldsymbol{a_{11}a_{22}a_{33}} - \boldsymbol{a_{11}a_{23}a_{32}} - \boldsymbol{a_{12}a_{21}a_{33}} + \boldsymbol{a_{12}a_{23}a_{31}} + \boldsymbol{a_{13}a_{21}a_{32}} - \boldsymbol{a_{13}a_{22}a_{31}}$,
si vede che ogni elemento è presente, ma con segno cambiato. Perché? Perché nelle permutazioni dei secondi indici (dello sviluppo di det(b)) abbiamo effettuato un'inversione di due secondi indici (due colonne), che ha cambiato il segno del prodotto.

In altre parole abbiamo la regola che, **scambiando fra loro due righe o due colonne di un determinante, il valore del terminante cambia segno.**

Ciò comporta che **un determinante che abbia due righe o due colonne eguali sia nullo,** perché, scambiando le due righe o colonne eguali da un lato deve restare lo stesso e dall'altro dovrebbe cambiare segno, il che lascia possibile solo il valore zero.

(Matrici e determinanti hanno anche un'interpretazione geometrica, per questo rimando al mio saggio sui determinanti
http://dainoequinoziale.it/sassolini/2019/11/19/determinanti.html, in questo stesso sito)

Possiamo ora immaginare di costruire una matrice costituita dai complementi algebrici, i cui elementi portano gli indici deigli elementi di cui sono il complemento, col segno come dalla figura 5. Una tale matrice porta diversi nomi. Quello di matrice "dei complementi", che indicheremo con M^* è uno di essi.

$$M^* = \begin{pmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{pmatrix}$$

Va però notato che nel calcolare il determinante, per esempio secondo gli elementi della prima riga, noi abbiamo moltiplicato la prima riga della M per la prima riga della M^*, cioè abbiamo eseguito un moltiplicazione riga per riga, non riga per colonna.
Il primo elemento della matrice prodotto N, che chiameremo P è infatti
$p_{11} = \det(M) = a_{11}A_{11} + a_{12}A_{12} + a_{13}A_{13}$ cioè $p_{11} = a_{1k}A_{1k}$, usando la nostra comoda convenzione di sommazione, estensibile a matrici quadrate di un numero qualsiasi di righe e colonne. Ricordo che per avere un prodotto "righe per colonne" dovremmo avere:

$$p_{11} = a_{1k}A_{k1}$$

che non funzionerebbe.
Dunque la matrice dei complementi M^*,

$$M^* = \begin{pmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{pmatrix}$$

non può essere utilizzata colle nostre regole per avere un determinante.
Per ritornare alla nostra moltiplicazione standard "righe per colonne", che dobbiamo mantenaere per coerenza, occorre trasformare la matrice M^* nella sua trasposta M^+, con righe e colonne scambiate, che chiameremo "matrice aggiunta".

$$M^+ = \begin{pmatrix} A_{11} & A_{21} & A_{31} \\ A_{12} & A_{22} & A_{32} \\ A_{13} & A_{23} & A_{33} \end{pmatrix}$$

E avremmo, ad esempio $p_{11} = \det(M) = a_{11}A_{11} + a_{12}A_{12} + a_{13}A_{13}$ dove $A_{ik} = A^T_{ki}$ della matrice trasposta, cioè $p_{11} = \det(M) = a_{1k}A^T_{k1} = a_{ik}A_{ik}$, sempre usando la comoda convenzione della somma sugli indici ripetuti. Notiamo che il determinante di M risulta dall'aver moltiplicato ***per esempio*** gli elementi di una riga ciascuno per i loro complementi. Ma che succederebbe se moltiplicassimo gli elementi di una riga/colonna per i complementi algebrici di un'altra riga/colonna? Non è una richiesta peregrina: per esempio, per trovare gli elementi p(1,2), p(1,3) dovremo moltiplicare ordinatamente gli elementi della riga 1 per i complementi che stanno nelle colonne 2, 3 etc., che sono però i complementi non di a(1,1) , a(1,2), a(1,3) ma di a(2,1) , a(2,2), a(2,3) e così via.
In altre parole, che succederebbe della somma $a_{ik}A_{lk}$ per i diverso da l ? Il fatto è che i complementi A_{lk} "sanno" che devono essere moltiplicati per i rispettivi a_{lk}. Se vengono moltiplicati per gli elementi a_{ik} (che già sono contenuti negli A(l,k),) si comportano come se il determinante avesse due righe o colonne eguali, e noi sappiamo che un tale determinante è nullo. Quindi deve essere $l = k$ e in conclusione , la moltiplicazione righe per colonne M x M^+ ci dà una matrice diagonale i cui elementi sono tutti eguali a det(M) e tutti gli altri sono nulli.

$$M\,M^{+} = \begin{pmatrix} detM & 0 & 0 \\ 0 & \det M & 0 \\ 0 & 0 & \det M \end{pmatrix}$$

Ora dividiamo ogni elemento di M^{+} per det M. Poiché ogni elemento non nullo, cioè ogni elemento diagonale, contiene una somma di addendi in ciascuno dei quali compare un solo elemento della matrice M^{+}, il det M può essere messo a fattore (o divisore) comune, e la matrice prodotto di M (M^{+}/detM) diventa, udite udite:

$$\text{M (M}^{+}\text{/detM)} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = I \quad la\ matrice\ identit\grave{a}.$$

In altre parole, abbiamo trovato la matrice inversa: $M^{-1} = M^{+}/$ (det M).

Ci sono altri modi di trovare la matrice inversa. In un determinante 2 x2 si tratta di risolvere quattro elementari equazioni di primo grado; se si è studiato di più, si può usare l'equazione di Cayley-Hamilton. Ma questo metodo mi sembra abbastanza interessante. Si noti bene, è fatto di passaggi semplici, facili da ricordare, ma non immediati da comprendere. Il passaggio che ho segnato in rosso è uno di questi: per capirlo bene bisogna pensarci un po'. D'altronde, si può anche fare un semplice esempio pratico con una matrice 3 x 3 , e si vede che effettivamente una colonna moltiplicata per i complementi algebrici di un'altra colonna dà risultato zero. Per sesempio, si supponga che nella

(A2) $$M = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

Siano a(1.1) = a(1,2); a(2,1)= a(2,2); a(3,1) = a(3,2). In tal caso la

(A3) $D = a_{11}a_{22}a_{33} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{31} - a_{13}a_{21}a_{31}$

Diventa

$D' = a_{11}a_{21}a_{33} - a_{11}a_{23}a_{31} - a_{11}a_{21}a_{33} + a_{11}a_{23}a_{31} + a_{13}a_{21}a_{31} - a_{13}a_{21}a_{31}$ =0

perché i termini si cancellano a due a due.

Torniamo ora alla soluzione dell'annunciato problema:

$$\boldsymbol{u} = \frac{1}{M}\,\boldsymbol{v} = M^{-1}\boldsymbol{v}$$

Occorre quindi

1) sostituire ogni elemento a_{ij} con il suo "cofattore" , il che ci dà la matrice M^{*};

2) scambiare le righe con le colonne della matrice così ottenuta, ottenendo una nuova matrice che chiameremo M^{+};

3) dividere ogni elemento di M^{+} per il determinante di M.

Si ottiene

(A3) $$\boldsymbol{u} = \frac{M^{+}}{\det(M)}\,\boldsymbol{v}$$

Questo è il modo più compatto di scrivere la regola. Ma il modo comunemente usato, dimostrabile independentemente, può essere estratto da questo:

Vediamo per esempio la componente z di **u** (non senza qualche attenzione agli indici e ai segni)**:**

$$z = u_3 = \frac{M^{+}_{3j}\,v_j}{\det M} = \frac{A_{31}v_1 + A_{32}v_2 + A_{33}v_3}{\det M} = \frac{\begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix} v_1 - \begin{vmatrix} a_{11} & a_{12} \\ a_{31} & a_{32} \end{vmatrix} v_2 + \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} v_3}{\det M}$$
$$= \frac{\begin{vmatrix} a_{11} & a_{12} & v_1 \\ a_{21} & a_{22} & v_2 \\ a_{31} & a_{32} & v_3 \end{vmatrix}}{\det M}$$

che corrisponde alle formule correnti della regola di Cramer (cfr. eq. 15).
La soluzione è unica perché è data dal prodotto della matrice inversa per il vettore **v**, che è ovviamente unico. Basta dunque dimostrare che anche la matrice inversa è unica. Infatti, se ci fosse una matrice N tale che $M\,N = M\,M^{-1} = I$, moltiplicando a sinistra per N, troveremmo che $N\,MN = N\,I = N$, e $N\,M\,M^{-1} = I\,M^{-1} = M^{-1}$. Cioè $N = M^{-1}$.
D'altronde, una soluzione non è possibile se il det(M) = 0.

Se ne deduce che:

A-I. Condizione necessaria e sufficiente perché una soluzione unica esista è che il determinante della matrice dei coefficienti sia diverso da zero.

Nel nostro caso

(A1) $$M = \begin{pmatrix} 1 & 4 \\ 1 & 1 \end{pmatrix}$$

il determinante vale D = -3, e quindi l'unica soluzione esiste.

> *Questa regola può essere estesa a n dimensioni. Se D proviene da un sistema di n equazioni, con n finito, l'equazione D = 0 certamente avrà i suoi n zeri, reali o complessi, singoli o multipli, che corrispondono agli autovalori del sistema. Tuttavia è bene notare che se n = ∞ (come nei casi che stiamo studiando), D può essere una funzione analitica priva di zeri nel campo reale, come exp(λ) nel qual caso ci dovremmo rassegnare a dire che il problema non possiede autovalori.*

Ora vogliamo modificare la nostra matrice in modo che trasformi il vettore originale **u**= (x, y) in se stesso, o almeno in **u** moltiplicato per un fattore ϱ (che chiameremo moltiplicatore, per non confonderlo con l'autovalore λ= 1/ϱ):

$$(A4) \qquad \begin{cases} x + 4\,y = \rho x \\ x + y = \rho y \end{cases}$$

$$(A5) \qquad \begin{cases} (1-\rho\,)x + 4\,y = 0 \\ x + (1-\rho\,)y = 0 \end{cases}$$

Ora, sostituendo x= 0 e y = 0, si vede che un tale sistema avrebbe l'ovvia soluzione **u** = (0, 0) = **0**. Ma questa sarebbe l'unica soluzione possibile, se il determinante dei coefficienti fosse diverso da zero. La sua utilità sarebbe zero, e quindi si chiama la "soluzione banale". Quindi , per avere anche soluzioni non banali, deve essere:

$$(A6) \qquad \begin{vmatrix} 1-\rho & 4 \\ 1 & 1-\rho \end{vmatrix} = 0$$

Questa equazione è di secondo grado ed ha due soluzioni:

$$\begin{cases} \rho_1 = 3 \\ \rho_2 = -1 \end{cases}$$

Sostituendo nella matrice vediamo che essa diventa:
Per ϱ= 3:

$$(A7) \qquad \begin{cases} -2x + 4\,y = 0 \\ x - 2y = 0 \end{cases}$$

Ove la seconda equazione è la prima moltiplicata per -1/2. Si può allora ricavare l'autovettore corrispondente al moltiplicatore ϱ= 3 scegliendo una qualunque delle due equazioni (A7), che ci dà un ***autovettore*** proporzionale a $\mathbf{u}^{(1)}$ = (2, 1). La costante di proporzionalità N può essere scelta tale N^2(2, 1)(2,1) = 1, cioè N = 1/ √5, il che ci dà un vettore normalizzato, cioè con lunghezza o norma = 1. Ma non ci serviremo di questa possibilità. Qui basta notare che qualsiasi vettore proporzionale all'autovettore trovato è ancora un autovettore.

Ponendo invece ϱ = -1 (l'altro autovalore), otteniamo:

$$(A8) \qquad \begin{cases} 2x + 4\,y = 0 \\ x + 2y = 0 \end{cases}$$

Ove di nuovo le due equazioni si sono ridotte a una, da cui otteniamo l'autovettore

$\mathbf{u}^{(2)}$ = (-2, 1), corrispondente al moltiplicatore ϱ= -1. La costante di proporzionalità sarà anch'essa M=N = 1/ √5.

A scopo pedagogico, facciamo il prodotto interno dei due vettori $\mathbf{u}^{(1)}$ = (2, 1) e $\mathbf{u}^{(2)}$ = (-2, 1), che vale (-2)2+ 1*1 = -3. *Se il risultato fosse nullo*, ricordando che il prodotto interno di due vettori a due dimensioni, cioè appartenenti a un piano, può essere scritto $\boldsymbol{u}^{(1)} \cdot \boldsymbol{u}^{(2)} = |\boldsymbol{u}^{(1)}|\,|\boldsymbol{u}^{(2)}|\cos(\theta)$, ciò significherebbe che, interpretando i nostri vettori come appartenenti a un piano, $\mathbf{u}^{(1)}$ e $\mathbf{u}^{(2)}$ sarebbero ortogonali, e potrebbero essere una base sostitutive di **i, j**. Ma evidentemente, qui ortogonali non lo sono.

Supponiamo ora di scegliere uno qualunque dei sistemi (A7) o (A8), ma con un termine noto (quasi) qualsiasi, per esempio (1, 3):
Ad esempio, dalla (A8), che corrisponde al moltiplicatore -1, abbiamo:

(A9)
$$\begin{cases} 2x + 4\,y = 1 \\ x + 2y = 3 \end{cases}$$

La soluzione di questo sistema, come si può verificare, non esiste. Per esempio, la regola di Cramer dà un denominatore, il determinante dei coefficienti, nullo. Se preferiamo, ricavando x dalle due equazioni ed eguagliando, si ottiene $3 - 2\,y = \left(\frac{1}{2}\right) - 2y$, da cui l'impossibile 3 = ½. Dunque ci si può chiedere se di soluzioni ne esistano.

Matrici trasposte e matrici simmetriche

La **matrice trasposta** T di M (che del resto abbiamo già incontrato) la si ottiene scambiando le righe colle colonne in modo ordinato, cioè la prima riga va nella prima colonna, la seconda riga nella seconda colonna eccetera, di modo che, chiamando M(i,j) gli elementi di M, dove il primo indice corrisponde alle righe, e il secondo alle colonne, avremo che T(i,j) = M(j,i), e, inevitabilmente, T(i, i)= M(i, i). Nel nostro caso:

$$T = \begin{pmatrix} 1 & 1 \\ 4 & 1 \end{pmatrix}$$

Il determinante di T è -3, eguale al determinante di M. Non è un caso, il determinante di una matrice trasposta è eguale al determinante della matrice originaria, come si verifica subita e come, con qualche fatica in meno, si dimostra.
Questo infatti deriva dalla definizione stessa di determinante. Come afferma il Tricomi (Lezioni di Analisi Matematica, Parte I):"*Noi chiameremo determinante il numero che risulta facendo la somma di tutti i prodotti che possono ottenersi moltiplicando fra loro ciascuna volta n quantità della tabella, scelte in righe e colonne diverse l'una dall'altra, essendo ciascun prodotto*

preso col proprio segno o col segno cambiato, secondo che le due permutazioni formate dai primi indici e dai secondi indici degli n fattori sono della stessa classe o di classi diverse".
Senza andare nel dettaglio dei significati dei vari termini che compaiono nella definizione, vediamo bene che in essa non si fa distinzione fra righe e colonne. I due termini potrebbero essere scambiati nell'unico luogo in cui appaiono, e nulla muterebbe della definizione. I primi indici sarebbero scambiati con i secondi indici, ma l'appartenenza alla stessa classe o a classi diverse resterebbe intatta. Credo che sia una delle dimostrazioni più brevi della matematica. L'unico punto necessario è che righe e colonne siano scambiate in modo ordinato, in modo che gli indici di ogni elemento non siano mutati, pur scambiandosi di posto.

A-II. Ne segue che anche l'equazione caratteristica della matrice trasposta (che deriva da un determinante in cui i moltiplicatori compaiono solo sulla diagonale principale), è la stessa, e quindi i moltiplicatori sono gli stessi.
Però, da questo non segue però che siano eguali anche gli autovettori. In effetti per la matrice T abbiamo:
(i) per il moltiplicatore $\varrho = 3$, l'autovettore (1, 2)
(ii) per il moltiplicatore $\varrho = -1$, l'autovettore (-1, 2)
Una **matrice simmetrica,** nella quale sono eguali gli elementi simmetrici (con indici scambiati) rispetto alla diagonale principale, è eguale alla sua matrice trasposta. La matrice simmetrica nel campo reale è il caso più elementare di una matrice ***autoaggiunta o hermitiana,*** nella quale gli elementi simmetrici (con indici scambiati) rispetto alla diagonale principale sono complessi coniugati l'uno dell'altro, e quelli della diagonale principale sono quindi reali.

Una proprietà delle **matrici simmetriche** è che gli autovettori sono ortogonali, cioè il loro prodotto interno è nullo. La dimostrazione non è difficile: Sia la matrice simmetrica

$$M = \begin{pmatrix} a & c \\ c & b \end{pmatrix}$$

Con due autovettori $a_1 = (u, v)$ e $a_2 = (w,z)$, corrispondenti a due diversi moltiplicatori/autovalori ϱ_1 e ϱ_2 rispettivamente, cioè tali che $M\, a_i = \varrho_i\, a_i$
Calcoliamo i due prodotti

$$\text{(A10)} \qquad (u,v)\begin{pmatrix} a & c \\ c & b \end{pmatrix}\begin{pmatrix} w \\ z \end{pmatrix} = (u,v)\begin{pmatrix} a\,w + c\,z \\ cw + bz \end{pmatrix} = uaw + ucz + vcw + vbz = \rho_1(uw + vz)$$

$$\text{(A11)} \qquad (w,z)\begin{pmatrix} a & c \\ c & b \end{pmatrix}\begin{pmatrix} u \\ v \end{pmatrix} = (w,z)\begin{pmatrix} a\,u + cv \\ c\,u + b\,v \end{pmatrix} = wau + vcw + zcu + zvb = \rho_2(uw + vz)$$

La differenza dà $(\rho_1 - \rho_2)(uw + vz) = (\rho_1 - \rho_2)(\boldsymbol{a_1} \cdot \boldsymbol{a_2}) = \mathbf{0}$.
Lo zero proviene dai quattro addendi $uaw + ucz + vcw + vbz$, manifestamente eguali a $wau + vcw + zcu + zvb$.
Poiché i due moltiplicatori sono diversi per ipotesi, ne segue che il prodotto interno $u\,w + vz$ dei due autovettori è nullo, cioòè i due autovettori sono ortogonali. Questo risultato si estende naturalmente alle matrici n x n, e anche alle matrici con n infinito e continuo (beninteso, se vi sono autovalori e quindi autovettori/autofunzioni.)

Scegliamo ora una matrice simmetrica elementare, ad esempio la

$$S = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$$

Si possono calcolare i due moltiplicatori, che risultano essere 2 e 0, e i relativi autovettori, (1,1) e (-1,1) rispettivamente, che, questa volta, a prima vista, sono ortogonali.

Ma la matrice trasposta porge un altro risultato interessante: torniamo a un'equazione del tipo (A9), in cui abbiamo inserito il moltiplicatore -1, che ha autovettore (-2, 1). La matrice trasposta, per lo stesso autovalore, porge l'autovettore (-1,2). Qui però vogliamo determinare un termine noto, un vettore (r,s) che renda una soluzione possibile:

(A10)
$$\begin{cases} 2x + 4\,y = r \\ x + 2y = s \end{cases}$$

Una soluzione sarà possibile solo se r ed s saranno tali per cui, a meno di un fattore di proporzionalità, le due equazioni divengono la stessa. Poiché la seconda equazione, *a parte il termine noto*, è eguale alla prima divisa per 2, occorre che anche r = 2s, cioè, in generale, r = 2k, s = k. Ma il vettore (2k, k) è ortogonale all'autovettore (-1, 2) della matrice trasposta, per lo stesso autovalore: infatti (-1)(2k) + k (2) = 0. Abbiamo dunque mostrato (non dimostrato) che

A-III. Se il determinante dei coefficienti è nullo, una soluzione è possibile purché il termine noto sia un vettore ortogonale all'autovettore della matrice trasposta corrispondente allo stesso moltiplicatore. Da cui deriva che se la matrice è simmetrica, basta che il vettore termine noto sia proporzionale a un altro autovettore (che è uno solo in una matrice 2x2), in quanto l'altro autovettore è ortogonale all'autovettore della matrice diretta, che è eguale alla matrice trasposta.

(Fine del Complemento A).

3. Estensione euristica del concetto di vettore.

Suppongo che sia noto al lettore che esistono almeno tre modi di trattare formalmente il concetto di vettore:

1) Vettore come "freccia", dotata di lunghezza, direzione, verso.
2) Vettore dato in termini delle sue componenti cartesiane, che altro non sono che le proiezioni del vettore sugli assi cartesiani. Ricordo che la proiezione del vettore non è altro che il "prodotto interno" del vettore sui versori degli assi cartesiani, che sono due (**i, j**) nel piano e tre (**i, j, k**) nello spazio. Un vettore può quindi essere rappresentato dalla semplice tabella delle sue componenti.
3) Vettore come matrice colonna: qui, le componenti possono essere disposte in una matrice colonna, alla quale si applica il calcolo matriciale.
4) Vettore come diagramma discreto. Si fissano sull'asse delle ascisse i punti 1,2,3, e in ordinata riportiamo il valore delle componenti 1,2,3 del vettore.

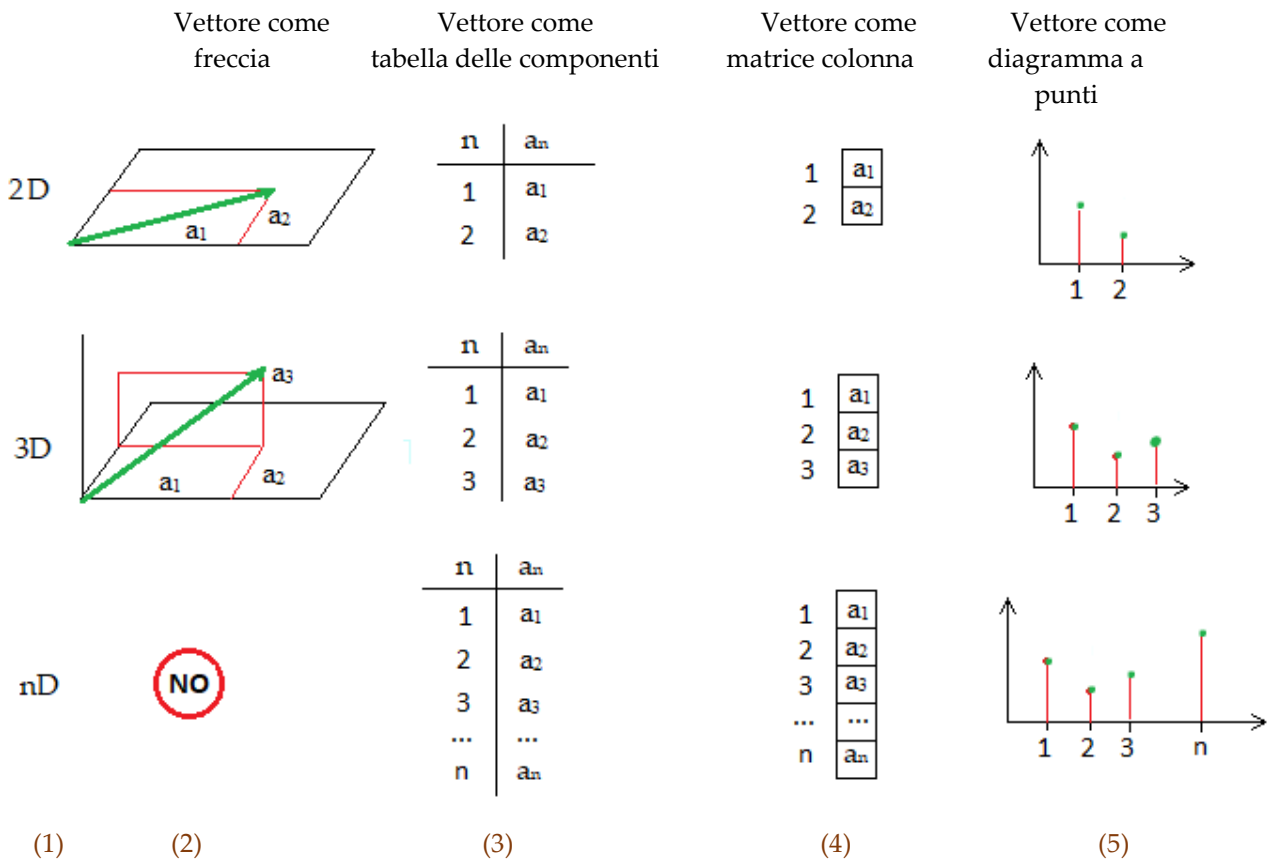


Fig.6

Qui dobbiamo introdurre un trucco, se non vogliamo addentrarci nel temuto campo degli spazi duali: un prodotto interno di due vettori rappresentati in forma matriciale (4) è dato dal prodotto di *un vettore riga (un vettore "trasposto")*, per un vettore colonna, seguendo le stesse regole del calcolo matriciale. Ne risulta un *numero*. Se si resta nel campo reale, il

prodotto non dipende dall'ordine dei due vettori moltiplicati: **u . v = v . u.** In effetti, nel campo complesso, le componenti del vettore "duale" dovrebbero essere le complesse coniugate del vettore originale.
Poiché, come è noto dal calcolo vettoriale, il prodotto interno o scalare di due vettori **u, v** è dato da:

$$\boldsymbol{u} \cdot \boldsymbol{v} = |\boldsymbol{u}||\boldsymbol{v}| \cos \vartheta$$

dove ϑ è l'angolo compreso fra i due vettori, $|\mathbf{u}|$ e $|\mathbf{v}|$ sono rispettivamente i moduli di **u** e **v**. Se il prodotto interno di due "vettori" non nulli è nullo, ciò significa che $\cos \vartheta = 0$, cioè $\vartheta = \frac{\pi}{2}$. Diremo allora che i due vettori sono "ortogonali".
La rappresentazione del vettore in termini di due versori **i, j ,** nel piano (bidimensionale) euclideo può essere estesa senza sforzo allo spazio tridimensionale euclideo, usando naturalmente tre versori **i, j, k** (colonna 2). Ma qui si fermano le possibilità di estensione.

La forma tabulare (colonna 3), invece, è ancora valida, anche se il vettore ha ***n*** componenti. Lo stesso vale per la forma matriciale (4). Infine, colonna (5), il vettore può essere rappresentato graficamente mettendo in fila le ***n*** componenti. Nelle forme (3), (4), (5) rappresentiamo, in altre parole, una funzione a(n) che ha valori solo per i valori interi di n, *Incidentalmente, non cambia nulla se invece di* a_n *scriviamo a(n), come abbiamo del resto già fatto.*

Fissiamoci un momento sulla rappresentazione (5) in Fig.6. Come abbiamo visto, il vettore è definito solo per determinati valori nei punti n dell'asse x, mentre in ordinate ogni componente può avere qualsiasi valore. Quando incominciamo a chiederci quale sia il valore dell'ordinata corrispondente a un numero non più intero, per esempio m = 1.3 (dove un'ordinata non è definita), ciò vuol dire che incominciamo a sentirci attratti dall'idea di "passare al continuo".

Fig. 5

Procedendo sempre euristicamente, possiamo dire allora che **una funzione è interpretabile come un vettore, con un'infinità continua di componenti.** Le ascisse fisse 1,2,3 saranno sostiuite da un continuo di ascisse che chiameremo x, una variabile reale che potrà assumere per la nostra esplorazione euristica tutti i valori reali, *tra i quali* gli interi 1,2,3 da cui siamo partiti.

E devo dire che, nei miei primi passi nella teoria delle funzioni, siccome nessuno mi aveva mostrato i semplici diagrammi di Fig.1 e Fig.2, non riuscivo a immaginarmi che cosa significasse la frase: **una funzione è interpretabile come un vettore che possiede un'infinità continua di componenti.** Invece si tratta di un'intuizione banale, che non riuscivo formarmi perché ero sul binario sbagliato:

Da dove partivo *da dove bisognava partire*

Fig.6

Tuttavia, si noti bene: anche se le componenti possono essere tradotte in una funzione f(x), questa funzione , nello spazio in cui ci troviamo, rappresenterà un unico punto o vettore.

In n dimensioni, esistono varie *matrici quadrate*, con n^2 elementi (n righe per n colonne), che si distinguono in due categorie: ***Operatori***, cioè matrici che rappresentano essenzialmente trasformazioni di vettori (quali rotazioni, allungamenti etc.) e ***Tensori***, *che rappresentano grandezze fisiche con* n^2 componenti, e quindi non rappresentabili con vettori. In certo senso anche queste possono agire su vettori trasformandoli, per cui la distinzione fra i due generi può diventare evanescente. *Noi, tuttavia ci accontenteremo di parlare di matrici che trasformano vettori*.

Come si è visto nel Complemento A, un operatore che non sia l'Identità, agendo su un vettore lo trasforma in un altro vettore. Tuttavia, dato un operatore K su uno spazio V, *possono esistere* certi vettori non nulli tali che K, operando su di essi, non ne cambi la "direzione" ma solo la "lunghezza", ovvero, *moltiplichi* il vettore per uno scalare.

(10) $$K\,\mathbf{v} = \lambda\,\mathbf{v}$$

Un tale vettore è detto "autovettore" (in inglese "eigenvector") e λ è il suo "autovalore (in inglese "eigenvalue"). Il simbolo λ, come vedremo, è sfortunato. In effetti, il mio Professore, F.G. Tricomi, nel suo corso di analisi, avrebbe scritto l'equazione (10) come

(10b) $$K\,\mathbf{v} = \varrho\,\mathbf{v}$$

e avrebbe chiamato ϱ col nome di "*moltiplicatore*", riservando il nome di "*autovalore*" al simbolo $\lambda = 1/\varrho$, che incontreremo. Il nome di moltiplicatore proviene evidentemente dal

fatto che l'operatore K agendo sul vettore **v** lo moltiplica per un fattore ϱ (cioè lo allunga o accorcia) senza cambiarne la direzione.

Se vale il prodotto righe per colonne, come noi abbiamo assunto e illustrato, la *lunghezza del vettore* è data dalla cosiddetta "norma" (o modulo) del vettore, considerato come matrice con una sola colonna. Si ottiene la norma facendo il prodotto delle componenti di un **vettore riga** opportuno (il vettore "trasposto" del vettore originario) per le corrispondenti componenti del nostro **vettore colonna** originale (terzo diagramma in Fig. 4) *ed estrendo la radice quadrata del risultato*. Perché la norma, *che è un* ***numero*** (terzo diagramma in Fig. 4) , abbia un valore reale positivo, occorre che le componenti corrispondenti del vettore riga siano le complesse coniugate di quelle del vettore colonna originale. Solo nel caso dei vettori che si incontrano nei primi due anni di Fisica, che hanno invariabilmente **componenti reali**, le componenti dei due vettori (originali e duali) sono eguali. Ma quando si procede nello studio, passando al campo complesso (ciò che è necessario in meccanica quantistica), la distinzione è necessaria. Questi vettori, riga e non colonna, con componenti complesse (come sarebbero le componenti dei vettori colonna con cui noi trattiamo usualmente), costituiscono lo "**spazio duale**", che nel nostro caso esiste sempre.

In questa prima esplorazione, ci accontenteremo di numeri reali.

Vediamo ora l'estensione (euristica) formale dei prodotti, restando nel campo dei numeri reali.

1) Prodotto interno.

(11) $$a \cdot b = \sum_{i=1}^{n} a(i)b(i)$$

Ma, considerando che per passare da a(i) b(i) a a(i+1)b(i+1) si può pensare di aver fatto un passo di lunghezza Δi = 1, possiamo senza molto sforzo immaginare di scrivere che

(11b) $$\boldsymbol{a} \cdot \boldsymbol{b} = \sum_{i=1}^{n} a(i)b\,(i)\,\Delta i \rightarrow \int_a^b a(x)b(x)dx$$

2) Prodotto di matrice per vettore colonna (il cui risultato, come si vede dal diagramma di Fermi, Fig.7) è un vettore colonna di n componenti (in quanto la matrice è quadrata). In questo contesto è immediato passare (intuitivamente) da un sistema di equazioni algebriche di primo grado a un'equazione integrale.

(12) $$\boldsymbol{K} \cdot \boldsymbol{v} = \sum_{j=1}^{n} K(i,j)v(j)\,\Delta j \rightarrow = \int_a^b K(x,y)v(y)dy$$

In effetti, un sistema di equazioni (per esempio di tre equazioni e tre incognite) può essere scritto come segue:

$$(13) \qquad \begin{cases} K(1,1)\,v(1) + K(1,2)\,v(2) + K(1,3)\,v(3) = u(1) \\ K(2,1)\,v(1) + K(2,2)\,v(2) + K(2,3)v(3) = u(2) \\ K(3,1)\,v(1) + K(3,2)\,v(2) + K(3,3)v(3) = u(3) \end{cases} \rightarrow \quad K \cdot \boldsymbol{v} = \boldsymbol{u}$$

E quindi, passando al continuo, possiamo dare un significato all'espressione:

$$(13\,b) \qquad \boldsymbol{K} \cdot \boldsymbol{v} = \sum_{i=1}^{n} K(i,j)v(j)\,\Delta j \rightarrow = \int_{\boldsymbol{a}}^{\boldsymbol{b}} \boldsymbol{K}(\boldsymbol{x},\boldsymbol{y})\boldsymbol{v}(\boldsymbol{y})\boldsymbol{dy} = \boldsymbol{u}(\boldsymbol{x})$$

Per ***uniformarci alla tradizione*** abbiamo dovuto mutare la notazione: la matrice K(i,j) è diventata una funzione in due variabili, K(x,y), dove x e y ora sostituiscono gli indici ***i*** e ***j***, il nome K proviene dal nome "nucleo", Kernel (tedesco o inglese) o kärna (svedese, la lingua di Fredholm, che per primo trattò le equazioni integrali sistematicamente); la sommatoria su j diventa un integrale su y, il vettore v(j) è una funzione v(y), della variabile y; il vettore u(i) è una funzione u(x), della variabile x.

Ma l'equazione (13 b) può essere interpretata in due modi differenti.

(i) Se il vettore **v** è il vettore incognito, e **u** è noto, allora abbiamo un sistema di equazioni che diventa **un'equazione integrale**, in cui la funzione incognita è la **v**. Provenendo (per analogia) da un sistema di equazioni lineari, si tratta di un'equazione integrale lineare, il cui nucleo K, funzione di x e y, moltiplica la funzione incognita v(y), la quale si trova sotto il segno di integrale, come si vede nella (13). **Un'equazione integrale non lineare (classe di funzioni che di qui innanzi ignoreremo)** viene scritta ad esempio come:

$$(14) \qquad \varphi(x) = f(x) + \lambda \int_a^b F(x,\xi,\varphi(\xi))d\xi$$

in cui la funzione $F\big(x,\xi,\varphi(\xi)\big)$ può assumere la forma più arbitraria possibile.

(ii) Se il vettore **v** è noto, allora la (13) indica una trasformazione di un vettore **v** in un nuovo vettore **u**. K potrebbe essere la matrice di rotazione, che, applicata a un vettore, ci dà le componenti del vettore ruotato. Passando all'integrale, abbiamo un'equazione formalmente identica alla (13), che però non viene più chiamata "equazione integrale", ma "**trasformata integrale**" di una funzione. Le due operazioni non sono lo stesso oggetto matematico, ma sono parenti stretti. Va detto che il campo delle trasformate integrali, tra le quali le più note sono quelle di **Fourier e di Laplace,** è un vasto campo della matematica, ricco di applicazioni pratiche, per esempio in elettrotecnica, del quale però non ci occuperemo.

Per un'equazione integrale, ha senso chiedersi a quale "classe funzionale" debbano appartenere il nucleo K(x,y) e il "termine noto" f(x) delle equazioni che studieremo. La risposta che utilizzeremo, valida per le equazioni "classiche" di Fredholm, è che K(x,y) sia una funzione "**al quadrato sommabile**", cioè appartenga allo spazio L^2, o spazio di Hilbert. Per lo spazio di Hilbert si veda eventualmente
http://dainoequinoziale.it/scienze/matematica/2020/02/03/hilbertspace.html

COMPLEMENTO B:

Sommario dei criteri perché K(x,y) appartenga allo spazio di Hilbert.

Diremo in breve che la funzione K(x,y) appartiene allo spazio di Hilbert se:

1) la funzione $K^2(x,y)$, considerata come funzione della sola x risulta **sommabile** (per noi sinonimo di **integrabile,** cioè tale che l'integrale esista e non sia infinito) per "quasi tutti" i valori di y. "Quasi tutti" significa che si possono accettare valori infiniti in punti che costituiscono un insieme "di misura nulla", intuitivamente un insieme di punti "isolati", come ad esempio un insieme numerabile, cioè tale che i suoi punti possano essere messi in corrispondenza biunivoca con i numeri naturali.

2) Lo stesso avviene scambiando le veci di x e y.

3) Le due funzioni positive (che per l'ipotesi (1) sono "quasi dappertutto esistenti")

$$(B1) \qquad A(x) = \sqrt{\int_a^b K^2(x,y)dy} \quad \text{e} \quad B(y) = \sqrt{\int_a^b K^2(x,y)dx}$$

appartengono alla classe L^2, cioè la norma sia "indotta dal prodotto interno", definito come:

$$(B2) \qquad f \cdot g = \int f^*(x)g(x)dx$$

Anche f(x), termine noto, deve appartenere alla classe L^2, o spazio di Hilbert.

Ammetteremo infine che valga il **teorema di inversione dell'ordine delle integrazioni,** che può essere così formulato:

Se esiste uno almeno dei due integrali ripetuti

$$(B3) \qquad \int_c^d dy \int_a^b dx\, F(x,y) \quad \text{e} \quad \int_a^b dx \int_c^d dy\, F(x,y)$$

allora esiste anche l'altro, ed è eguale al primo.

COMPLEMENTO C:

Perché vogliamo che K(x,y) e f(x) appartengano allo spazio di Hilbert, L^2 ?

La categoria dei cosiddetti spazi L^p (chiamati anche spazi di **Lebesgue**, forse a torto) è un'importante classe di spazi di **Banach** (per cui si veda il mio saggio sullo spazio di Hilbert sopracitato). Uno spazio di Banach, in poche parole, ha la

caratteristica di essere uno spazio ***topologico, vettoriale (o lineare), metrico, normato, completo.***

E, precisa it.wikipedia:

Si definisce norma p-esima o norma L^p di f il numero

$$\|f\|_p = \left(\int_X |f|^p d\mu\right)^{\frac{1}{p}}$$

Lo spazio delle funzioni misurabili con norma L^p finita è detto $L^p(X, \mathfrak{M}, \mu)$, o anche $L^p(X)$, $L^p(\mu)$ o solo L^p. Le funzioni in L^p si dicono *a p-esima potenza sommabile.*

Ma perché, tra tutti gli spazi L^p, vogliamo che K e f appartengano allo spazio in cui p=2, cioè allo spazio L^2? Perché lo spazio L^2 è, tra gli spazi L^p, l'unico che gode della proprietà di derivare la sua norma dal prodotto interno nella forma (A2) qui sopra, e quindi è l'unico spazio di Hilbert. Il vantaggio della struttura del prodotto interno nello spazio di Hilbert è che esso permette la più semplice estensione dei concetti dello spazio euclideo agli spazi di funzioni, e quindi l'uso di vettori/funzioni ortogonali, e tutto ciò che ne segue, con applicazioni che vanno dalle serie di Fourier alla meccanica quantistica. Faremo abbondante uso di queste estensioni intuitive (Complemento A e parte 4) in questo lavoro.

Ho volutamente omesso di dire nelle precedenti condizioni di appartenenza alla classe L^2 che i vari integrali sono oggi generalmente intesi "nel senso di Lebesgue". Ora, Lebesgue introdusse il suo concetto di integrale nel 1902, mentre Fredholm pubblicava la soluzione delle sue equazioni nel 1903. Dal suo testo pare che Fredholm non cercasse la generalità offerta dall'integrazione secondo Lebesgue (che non è mai menzionato nei suoi articoli fondamentali), e si servisse dell'integrazione comunemente nota ai suoi tempi, nel senso di Riemann, che egli usa senza neppure menzionare l'autore. Poiché noi vogliamo soltanto dare un'idea non rigorosa della soluzione data da Fredholm, e d'altra parte questi non fece ricorso all'integrale di Lebesgue, neanche noi parleremo di integrali nel senso di Lebesgue, che potrebbero solo complicare una presentazione elementare come questa.

4. La ricerca degli autovalori di una matrice. Analogie.

Diventa immediato scrivere un'equazione agli autovalori, che, come ricordiamo, è:

(10) $$K\,\mathbf{v} = \lambda\,\mathbf{v}$$

Ma Fredholm diede il nome di λ all'inverso del λ dell'equazione (10). In altre parole, la sua equazione di partenza fu

(10c) $$\mathbf{v} = \lambda\,K\,\mathbf{v}$$

Si vede così che, per chiarezza, aveva un senso chiamare *moltiplicatore* il valore λ nell'equazione (10), e assegnargli un simbolo diverso, cioè $\varrho = 1/\lambda$. Ci sono testi , anche autorevoli, che si infischiano allegramente di questa ambiguità.

Ora, io non so se Fredholm abbia seguito, anche solo come primo abbozzo, il ragionamento che consiste nel partire da un sistema di equazioni algebriche, per giungere alla soluzione dell'Everest delle equazioni integrali lineari, cioè alla sua equazione generale (**Equazione di Fredholm di seconda specie**), con nucleo qualsiasi, soggetto a molto liberali condizioni (descritte nel Complemento A):

(14) $$\varphi(x) = f(x) + \lambda \int_a^b K(x,\xi)\varphi(\xi)d\xi$$

ove φ(x) è la funzione cercata e f(x) è una funzione nota. In forma vettore-operatore, la funzione cercata corrisponderebbe più precisamente al vettore **v** nell'equazione:

(14b) $$\mathbf{v} = \boldsymbol{u} + \lambda \mathrm{K}\, \mathbf{v}$$

Ovvero

(14c) $$(1 - \lambda K)\boldsymbol{v} = \boldsymbol{u}$$

Si può trarre uno sviluppo in forma di serie geometrica, che non semplifica necessariamente il nostro compito.

$$\boldsymbol{v} = \frac{1}{1-\lambda K}\boldsymbol{u} = (1 + \lambda K + \lambda^2 K^2 \dots)\boldsymbol{u} = \boldsymbol{u} + \lambda K[\lambda K + \lambda^2 K^2 \dots]\boldsymbol{u}$$

In termini di vettori e matrici, l'equazione (14c), per vettori a tre componenti, diventa la

(14d) $$\begin{cases} \big(1-\lambda K(1,1)\big)\varphi(1) - \lambda K(1,2)\varphi(2) - \lambda K(1,3)\varphi(3) = f(1) \\ -\lambda K(2,1)\varphi(1) + (1-\lambda K(2,2))\varphi(2) - \lambda K(2,3)\varphi(3) = f(2) \\ -\lambda K(3,1)\varphi(1) - \lambda K(3,2)\varphi(2) + (1-\lambda K(3,2))\varphi(3) = f(3) \end{cases}$$

Qui il vettore ***v, ignoto,*** ha componenti $(\varphi(1), \varphi(2), \varphi(3))$, e il vettore ***u, noto,*** ha componenti *(f(1), f(2), f(3))*. Tanto nel nucleo quanto nei vettori, i numeri 1,2,3 stanno ora per x_1, x_2, x_3, e descrivono i punti di una funzione, a distanza 0.5 uno dall'altro, nell'intervallo (0,1). Ma nulla ci impedisce di aumentare il numero di punti (e di equazioni) diminuendo la distanza tra i punti a valori δ, e finalmente a valori dx infinitesimi.

Questa scelta di nomi, come vedremo, corrisponde all'approccio di Fredholm, che probabilmente pensò (tacitamente) di passare dalla (14c) alla (14d) e poi di estendere il sistema che compare nella (14d) ad un sistema di infinite componenti discrete, in cui gli intervalli che separano i punti punti 1 da 2 e 2 da 3 diminuiscono fino a diventare infinitesimi e permettere di passare al continuo.

Secondo la regola di Cramer, la soluzione di (14d) è data da:

$$(15)\qquad \varphi(1)=\frac{\begin{matrix} f(1) & -\lambda K12 & -\lambda K13 \\ f(2) & 1-\lambda K22 & -\lambda K23 \\ f(3) & -\lambda K32 & 1-\lambda K33 \end{matrix}}{\begin{matrix} 1-\lambda K11 & -\lambda K12 & -\lambda K13 \\ -\lambda K21 & 1-\lambda K22 & -\lambda K23 \\ -\lambda K31 & -\lambda K32 & 1-\lambda K33 \end{matrix}};\quad \varphi(2)=\frac{\begin{matrix} 1-\lambda K11 & f(1) & -\lambda K13 \\ -\lambda K21 & f(2) & -\lambda K23 \\ -\lambda K31 & f(3) & 1-\lambda K33 \end{matrix}}{\begin{matrix} 1-\lambda K11 & -\lambda K12 & -\lambda K13 \\ -\lambda K21 & 1-\lambda K22 & -\lambda K23 \\ -\lambda K31 & -\lambda K32 & 1-\lambda K33 \end{matrix}};\quad \varphi(3)=\frac{\begin{matrix} 1-\lambda K11 & -\lambda K12 & f(1) \\ -\lambda K21 & 1-\lambda K22 & f(2) \\ -\lambda K31 & -\lambda K32 & f(3) \end{matrix}}{\begin{matrix} 1-\lambda K11 & -\lambda K12 & -\lambda K13 \\ -\lambda K21 & 1-\lambda K22 & -\lambda K23 \\ -\lambda K31 & -\lambda K32 & 1-\lambda K33 \end{matrix}}$$

che si può riassumere come:

$$(16)\qquad \varphi(i)=\frac{\sum_j f(j)D(i,j;\lambda)}{D(\lambda)} \quad\rightarrow\quad \varphi(x)=\int_a^b \frac{D(x,\xi;\lambda)}{D(\lambda)} f(\xi)d\xi$$

in cui D(λ) è il determinante dei coefficienti e D(i,j; λ) è il "complemento algebrico" di f(j).

Si noti che in uno sviluppo in serie di C, i soli termini in cui il parametro λ non appare in D(i,j; λ) sono rispettivamente:
(i) per $\varphi(1)$, *il termine* f(1);
(ii) per $\varphi(2)$, *il termine* f(2) ;
e (iii) per $\varphi(3)$, *il termine* f(3)
- in una parola, per $\varphi(i)$, f(i). Si tratta dei soli elementi, in rosso nella (15), che sono moltiplicati per i due addendi 1 e possono essere estratti dall'integrale, in quanto non si integra su x, che è la variabile indipendente della funzione φ.

A questo punto il **Whittaker and Watson, p. 214**, fa un'osservazione che mi lascia poco convinto, anche se in un modo o nell'altro conduce al risultato che vogliamo:
"It is easily seen that the appropriate limiting form to be associated to D(i,j; λ) for i= j is D(λ)" (si vede facilmente che la forma limite appropriata da associarsi a D(i,j; λ) per i = j è D(λ)). Per esempio, dalla prima delle (15), noi abbiamo

$$\varphi(1)=\frac{\begin{matrix} f(1) & -\lambda K12 & -\lambda K13 \\ f(2) & 1-\lambda K22 & -\lambda K23 \\ f(3) & -\lambda K32 & 1-\lambda K33 \end{matrix}}{\begin{matrix} 1-\lambda K11 & -\lambda K12 & -\lambda K13 \\ -\lambda K21 & 1-\lambda K22 & -\lambda K23 \\ -\lambda K31 & -\lambda K32 & 1-\lambda K33 \end{matrix}}=\frac{f(1)[\ 1-\ \lambda K22-\lambda K33+\lambda^2 K22\,k33-\lambda^2 K23\,K33]+altri\ termini}{\begin{matrix} 1-\lambda K11 & -\lambda K12 & -\lambda K13 \\ -\lambda K21 & 1-\lambda K22 & -\lambda K23 \\ -\lambda K31 & -\lambda K32 & 1-\lambda K33 \end{matrix}}$$

che, al limite del continuo, dovrebbe diventare $\varphi(x)=\int(f(x)\,D(\lambda)\ +\ altri\ termini)/\ D(\lambda))$.

Più correttamente, facendo uscire f(x) dall'integrale, che agisce solo sulla variabile ξ, e $D(\lambda)$, che non contiene variabili, dovremmo poter scrivere:

$$(17)\qquad \varphi(x)=f(x)+\int_a^b \frac{D(x,\xi;\lambda)}{D(\lambda)} f(\xi)d\xi$$

e dall'integrale sarebbero esclusi i termin $D(x,x;\lambda)$.

Che così stiano realmente le cose, può essere rigorosamente dimostrato (Fredholm, e.i (1903): "Sur une classe d'equations fonctionnelles" *Acta Mathematica*. **27**: 365–390.) Tuttavia, dire che sono perfettamente convinto del suggerimento di Whittaker e Watson, sarebbe un'esagerazione.

Indubbiamente resta una zona di nebbia. Ma è difficile ottenere altrimenti la (17), e d'altra parte si può pensare che l'espressione per $\varphi(2)$, avvicinandosi il punto 2 a distanza infinitesima dal punto 1, ci possa restituire in certo modo la prima riga e la prima colonna, senza necessariamente privarci della seconda riga e della seconda colonna, che però già avevamo, e riavremo da $\varphi(3)$, e così via.

È dunque relativamente facile scrivere per analogia la (17), tanto più se sappiamo dove vogliamo arrivare, ma il problema è trovare l'espressione del rapporto $\frac{D(x,\xi;\lambda)}{D(\lambda)}$. La soluzione trovata *rigorosamente* da Fredholm (sudando matematico sangue) fu effettivamente la (17), in cui il rapporto $\frac{D(x,\xi;\lambda)}{D(\lambda)}$ è detto "nucleo risolvente", e le due funzioni $D(x,\xi;\lambda)$ *e* $D(\lambda)$ sono due estensioni del concetto di determinante ad una matrice *infinita e continua* (!), che vedremo più avanti.

Si noti che $D(\lambda)$ può essere portato fuori dall'integrale, in quanto non dipende dalle variabili x e ξ. Io ho preferito lasciarlo nell'integrale per sottolineare la somiglianza formale colla soluzione di Cramer (15) dei sistemi di equazioni algebriche lineari.

Critica per la soluzione è la condizione che $D(\lambda)$ non sia nullo. I valori di λ per cui ciò avviene sono, come ci si può aspettare, gli autovalori del problema, in certo senso la parte più interessante dell'equazione. Detto brutalmente, ci interessa soprattutto sapere a quali velocità l'asse dell'elica rischia di spaccarsi, non la forma che l'asse assume nel corso della rotazione.

Tuttavia, mentre io preferisco seguire il cammino dalle matrici finite alle equazioni integrali, Fredholm ***pubblicò*** il cammino inverso, dalle equazioni integrali alle matrici, spezzettando le funzioni φ(x) ed f(x) nell'intervallo (a,b), **per noi (0,1),** in un numero n (finito) di segmenti eguali (x_i ,x_{i+1}), e il nucleo K(x,y) in una matrice n x n, corrispondente agli stessi segmenti Egli trasformò quindi l'equazione integrale in un sistema di equazioni algebriche, da risolversi con un sistema non molto diverso concettualmente dalla regola di Cramer, per poi passare al limite del continuo. Ma occorre dire che il mio insegnante di Istituzioni di Analisi Superiore, il Prof. F.G. Tricomi, che non si spaventava certo davanti ai calcoli, e scrisse un testo specialistico proprio sulle Equazioni Integrali, tuttavia nel suo corso ***"Istituzioni di Analisi Superiore" (Gheroni, 1962),*** arrivò a scrivere: "*Fredholm ha dato degli espliciti formalmente molto eleganti sviluppi in serie di una possibile coppia di queste funzioni* $D(x,\xi;\lambda)$ *e* $D(\lambda)$*, che però, in pratica, si sono rivelati di scarsa utilità a causa della loro troppo grande complicazione sostanziale*". Dopodiché, nel testo di "Istituzioni", F.G. Tricomi procede col trattare casi particolari e semplificati del nucleo K(x,y), quali i casi studiati da **Volterra, da Pincherle e Goursat** etc., che esulano da questa nostra prima esplorazione.

Va detto che le equazioni integrali di Fredholm di seconda specie tendono a non essere più facili da risolvere "in forma chiusa" delle equazioni differenziali, ma permettono una soluzione numerica approssimata, proprio spezzettando la funzione in segmenti e

trasformando l'equazione integrale in un sistema di equazioni algebriche. Talvolta occorre accontentarsi del risultato numerico approssimato, talaltra passando al limite si ottiene una funzione continua nota (ma ciò avviene di rado e solitamente solo sui libri di esercizi). Ci sono anche casi, ad esempio le equazioni dette di **Volterra**, in cui il limite superiore dell'integrale è x, ovvero il nucleo è nullo per y > x, che non di rado sono relativamente semplici da risolvere.

Il compito che mi propongo ora, è quello di proporre un procedimento intuitivo per arrivare ai "formalmente molto eleganti sviluppi in serie" citati dal Tricomi, e quindi a un'idea della (quasi inutilizzabile) soluzione generale di Fredholm per la più generale equazione. E così, finalmente potrò togliermi dalla scarpa il sassolino costituito dalla frase del Tricomi in corsivo, citata più sopra, che forse , tra tutti gli allievi del corso di allora, diede fastidio solo a me.

5. I determinanti di Fredholm.

Capire come ci arrivò Fredholm, non è banale, e qui io sto certamente facendo della fanta-matematica. Ma non saprei che altro fare.

Anzitutto, penso per motivi estetici e di tradizione, Fredholm voleva assai probabiblmente riprodurre fin da principio la **regola di Cramer**, che produce la soluzione in forma di rapporto di due determinanti. Penso che questa sia una parte non pubblicata della sua ricerca.

Il suo problema era quindi quello di sviluppare **due determinanti infiniti e continui.** Sfortunatamente, un determinante di una matrice infinita e continua, che io sappia, non può essere calcolato con i metodi comunemente utilizzati per calcolare determinanti finiti, come la regola di Laplace (pag.18), anzi, il suo calcolo non può neppure incominciare dai metodi solitamente usati.

5.1: Il determinante a denominatore, $\boldsymbol{D(\lambda)}$.

Incominciamo con il determinante più semplice dei due, $D(\lambda)$.

La prima idea era quella di svilupparlo in serie delle potenze di un parametro, che si offriva spontaneamente, ed era (il suo) λ. Si vede subito, che se λ non fosse stato portato a fattore della matrice, tenere conto dei vari termini sarebbe stato più complicato. Non era naturalmente garantito che questo parametro sarebbe stato piccolo, né che la serie sarebbe stata una serie convergente.

I passi che si offrivano a Fredholm erano quindi probabilmente:

1) sviluppare come esperimento un determinante FINITO in termini delle potenze del parametro λ. Trovò che il determinante finito poteva essere scritto come un polinomio in potenze di λ, i cui coefficienti erano a loro volta dei determinanti estratti dal determinante completo dei coefficienti.

2) Si trattava ora di vedere se i singoli determinanti finiti che apparivano come coefficienti delle varie potenze di λ presentassero regolarità formali che permettessero di estenderli ad un determinante infinito;

3) Passare al limite del continuo. Qui Fredholm non aveva molta scelta, se voleva una semplice soluzione. In un modo o nell'altro, *doveva ottenere delle semplici sommatorie, magari a più indici,* che avrebbero potuto facilmente trasformarsi in integrali a più variabili di integrazione, come abbiamo euristicamente mostrato più sopra. È infatti semplice estendere euristicamente le sommatorie ad integrali, se ci si può servire dell'analogia:

(18) $$\sum_{i,j,k\ldots=0\ o1}^{n} F(i,j,k\ldots.)\Delta i\ \Delta j\ \Delta k\ldots \rightarrow \iiint_{0\ o\ 1}^{n} F(x,y,z\ldots)dz\,dy\,dz\ldots$$

Tuttavia, $D(x,\xi;\lambda)$ è più complicato da scriversi nella forma del membro di sinistra, in quanto bisogna introdurre la funzione *f(x)* nella colonna giusta del determinante.

I. D(λ) : Determinante finito.

Come primo esempio, scegliamo un determinante 3 x 3, con l'autovalore (λ di Fredholm) che moltiplica la matrice. La matrice sarà denominata K, e gli elementi di matrice K11, K12, K13 etc. Qui, come è ovvio, K11 sta per K(1,1) o anche K_{11} e via dicendo.

Vogliamo sviluppare il determinante: (19)

$$\begin{vmatrix} 1-\lambda K11 & -\lambda K12 & -\lambda K13 \\ -\lambda K21 & 1-\lambda K22 & -\lambda K23 \\ -\lambda K31 & -\lambda K32 & 1-\lambda K33 \end{vmatrix}$$

Sviluppando il determinante come si preferisce e ordinando secondo le potenze di λ si ottiene:

(20) $1-\lambda\,(K11+K22+K33\,)+\lambda^2(-K12\ \ K21+K11\,K22-K13\,K31-K23\,K32+K11K33+K22K33)-\lambda^3\,(-K13\,K22\,K31+K12\,K23\,K31+K13\,K21\,K32-K11\,K23\,K32-K12\,K21\,K33+K11\,K22\,K33)$

Il primo termine, 1, coefficiente di λ^0, proviene dal prodotto dei termini 1 della diagonale principale. Si può facilmente comprendere che, se anche la matrice diventasse infinita, il termine non cambierebbe.

Il secondo termine, il coefficiente di λ^1, proviene dalla somma dei secondi termini degli elementi diagonali, tutti negativi, ciascuno moltiplicato per tutti i primi termini (tutti eguali a 1); altri termini non possono entrare nel prodotto , in quanto introdurrebbero un fattore λ , e quindi andrebbero di diritto a far parte di un termine di ordine superiore a λ^1.

Questo secondo termine, che potrebbe essere scritto come $-\lambda \sum_1^3 K(i,i)\Delta i$, può facilmente passare al continuo, diventando $-\lambda \int_0^1 K(x,x,)dx$.

Più interessante è il terzo termine in λ^2, che, grazie al prodotto di due fattori -λ, è positivo. Il termine era destinato a contenere tutti i determinanti 2 x 2. Fredholm potrebbe aver pensato: "Noi già abbiamo calcolato *direttamente* il valore del coefficiente di λ^2. Calcoliamo un sommatorio di tutti i determinanti 2 x2 possibili nel nostro determinante 3 x 3 e vediamo se ci sono delle regolarità interessanti."

Il valore calcolato direttamente è dato da sei termini, che sembra si possano ridurre a coppie, una positiva e una negativa, come termini di tre determinanti.

(21) $$(-\mathrm{K12}\ \mathrm{K21} + \mathrm{K11}\ \mathrm{K22} - \mathrm{K13}\ \mathrm{K31} + \mathrm{K11}\ \mathrm{K33} - \mathrm{K23}\ \mathrm{K32} + \mathrm{K22}\ \mathrm{K33})$$

Notiamo che se si tratta dei risultati di tre determinanti, essi hanno i termini K11, K22, K33 nella diagonale principale. Mettendo insieme le coppie di termini che hanno gli stessi due indici (1 e 2; 1 e 3; 2 e 3), vediamo che essi possono risultare dalla sommatoria:

$$\sum_{i,j=1}^{3} \begin{vmatrix} K(i,i) & K(i,j) \\ K(j,i) & K(j,j) \end{vmatrix}$$

La sommatoria produce i seguenti NOVE determinanti:

$$(1) \begin{vmatrix} K11 & K11 \\ K11 & K11 \end{vmatrix}; (2) \begin{vmatrix} K22 & K22 \\ K22 & K22 \end{vmatrix}; (3) \begin{vmatrix} K33 & K33 \\ K33 & K33 \end{vmatrix};$$

$$(4) \begin{vmatrix} K11 & K12 \\ K21 & K22 \end{vmatrix}; (5) \begin{vmatrix} K11 & K13 \\ K31 & K33 \end{vmatrix}; (6) \begin{vmatrix} K22 & K23 \\ K32 & K33 \end{vmatrix};$$

$$(7) \begin{vmatrix} K22 & K21 \\ K12 & K11 \end{vmatrix}; (8) \begin{vmatrix} K33 & K31 \\ K13 & K11 \end{vmatrix}; (9) \begin{vmatrix} K33 & K32 \\ K23 & K22 \end{vmatrix};$$

Ma guardiamo meglio: i determinanti (1),(2),(3) sono tutti nulli, perché hanno tutti gli elementi eguali; mentre le coppie di determinanti (4) e (7), (5) e (8), (6) e (9) sono eguali a due a due, avendo righe e colonne scambiate. In altre parole, il risultato della sommatoria è

$$\sum_{i,j=1}^{3} \begin{vmatrix} K(i,i) & K(i,j) \\ K(j,i) & K(j,j) \end{vmatrix} = 2\ (+\mathrm{K11}\ \mathrm{K22} - \mathrm{K12}\ \mathrm{K21} + \mathrm{K11}\ \mathrm{K33} - \mathrm{K13}\ \mathrm{K31} + \mathrm{K22}\ \mathrm{K33} - \mathrm{K23}\ \mathrm{K32})$$

Il secondo termine, che noi avevamo calcolato direttamente, è quindi la metà della sommatoria dei determinanti. Possiamo dunque euristicamente passare dal determinante

3x3 al determinante n x n, e poi al determinante per n → ∞ , e quindi al continuo, introducendo i soliti segmenti Δi = Δj =1. Identificando 1 con x e 2 con y otteniamo i nostri primi tre termini:

(22) $$D(\lambda) = 1 - \lambda \int_0^1 K(x,x,)dx + \frac{1}{2}\lambda^2 \iint_0^1 \begin{vmatrix} K(x,x) & K(x,y) \\ K(y,x) & K(y,y) \end{vmatrix} dx\, dy - \cdots$$

Vorrrei ora calcolare almeno il quarto termine, che moltiplica $(-\lambda)^3$, negativo, il cui valore è già stato direttamente calcolato ed è:

(23) $D = (+K11\ K22\ K33 + K12\ K23\ K31 + K13\ K21\ K32 - K11\ K23\ K32 - K12\ K21\ K33 - K13\ K22\ K31)$

Una tripla sommatoria da 1 a 3 ci darebbe 27 determinanti 3 x 3, e ciascuno di essi sei termini.

Ecco, nella forma prediletta da MATHEMATICA, i primi nove di queste ventisette determinanti.

(1) $\begin{vmatrix} K[1,1] & K[1,1] & K[1,1] \\ K[1,1] & K[1,1] & K[1,1] \\ K[1,1] & K[1,1] & K[1,1] \end{vmatrix}$; (2) $\begin{vmatrix} K[1,1] & K[1,1] & K[1,2] \\ K[1,1] & K[1,1] & K[1,2] \\ K[2,1] & K[2,1] & K[2,2] \end{vmatrix}$; (3) $\begin{vmatrix} K[1,1] & K[1,1] & K[1,3] \\ K[1,1] & K[1,1] & K[1,3] \\ K[3,1] & K[3,1] & K[3,3] \end{vmatrix}$

(4) $\begin{vmatrix} K[1,1] & K[1,2] & K[1,1] \\ K[2,1] & K[2,2] & K[2,1] \\ K[1,1] & K[1,2] & K[1,1] \end{vmatrix}$; (5) $\begin{vmatrix} K[1,1] & K[1,2] & K[1,2] \\ K[2,1] & K[2,2] & K[2,2] \\ K[2,1] & K[2,2] & K[2,2] \end{vmatrix}$; (6) $\begin{vmatrix} K[1,1] & K[1,2] & K[1,3] \\ K[2,1] & K[2,2] & K[2,3] \\ K[3,1] & K[3,2] & K[3,3] \end{vmatrix}$

(7) $\begin{vmatrix} K[1,1] & K[1,3] & K[1,1] \\ K[3,1] & K[3,3] & K[3,1] \\ K[1,1] & K[1,3] & K[1,1] \end{vmatrix}$; (8) $\begin{vmatrix} K[1,1] & K[1,3] & K[1,2] \\ K[3,1] & K[3,3] & K[3,2] \\ K[2,1] & K[2,3] & K[2,2] \end{vmatrix}$; (9) $\begin{vmatrix} K[1,1] & K[1,3] & K[1,3] \\ K[3,1] & K[3,3] & K[3,3] \\ K[3,1] & K[3,3] & K[3,3] \end{vmatrix}$

Vediamo subito che solo qualcuno di questi determinanti si salva:
(1) Ha tutte le righe e le colonne eguali, e quindi è nullo;
(2), (3) hanno la prima e la seconda colonna eguali, e quindi sono nulli;
(4) (7) hanno la prima e la terza colonna eguali, e quindi sono nulli;
(5) ha la seconda e la terza colonna eguali e quindi è nullo;
(9) ha due righe e due colonne eguali e quindi è nullo;
(6) e (8) non sono a priori nulli, ma si riducono allo stesso determinante scambiando la seconda e la terza riga;
D' = K11 K22 K33 + K12 K23 K31+ K13 K21 K32 - K11 K23 K32 – K12 K21 K33 –K13 K22 K31
Da paragonarsi con:
D = K11 K22 K33 + K12 K23 K31 + K13 K21 K32 − K11 K23 K32 − K12 K21 K33 − K13 K22 K31,
che è il determinante che abbiamo ricavato direttamente.
Quindi il risultato della sommatoria di questi primi nove determinanti è 2D.

E' ora il turno dei secondi nove determinanti:

$$(10)\begin{vmatrix} K[2,2] & K[2,1] & K[2,1] \\ K[1,2] & K[1,1] & K[1,1] \\ K[1,2] & K[1,1] & K[1,1] \end{vmatrix};\ (11)\begin{vmatrix} K[2,2] & K[2,1] & K[2,2] \\ K[1,2] & K[1,1] & K[1,2] \\ K[2,1] & K[2,2] & K[2,2] \end{vmatrix};\ (12)\begin{vmatrix} K[2,2] & K[2,1] & K[2,3] \\ K[1,2] & K[1,1] & K[1,3] \\ K[3,2] & K[3,1] & K[3,3] \end{vmatrix}$$

$$(13)\begin{vmatrix} K[2,2] & K[2,2] & K[2,1] \\ K[2,2] & K[2,2] & K[2,1] \\ K[1,2] & K[1,2] & K[1,1] \end{vmatrix};\ (14)\begin{vmatrix} K[2,2] & K[2,2] & K[2,2] \\ K[2,2] & K[2,2] & K[2,2] \\ K[2,2] & K[2,2] & K[2,2] \end{vmatrix};\ (15)\begin{vmatrix} K[2,2] & K[2,2] & K[2,3] \\ K[2,2] & K[2,2] & K[2,3] \\ K[3,2] & K[3,2] & K[3,3] \end{vmatrix}$$

$$(16)\begin{vmatrix} K[2,2] & K[2,3] & K[2,1] \\ K[3,2] & K[3,3] & K[3,1] \\ K[1,2] & K[1,3] & K[1,1] \end{vmatrix};\ (17)\begin{vmatrix} K[2,2] & K[2,3] & K[2,2] \\ K[3,2] & K[3,3] & K[3,2] \\ K[2,2] & K[2,3] & K[2,2] \end{vmatrix};\ (18)\begin{vmatrix} K[2,2] & K[2,3] & K[2,3] \\ K[3,2] & K[3,3] & K[3,3] \\ K[3,2] & K[3,3] & K[3,3] \end{vmatrix}$$

In questo gruppo di nove, abbiamo che

I seguenti determinanti sono nulli:

- Tutti gli elementi eguali in (14)
- Due colonne righe e due colonne eguali in (10), (13), (15), (17), (18)
- Due colonne eguali in (11)

Invece, due determinanti sono eguali in (12) e (16), con valore

$$-K13\ K22\ K31 + K12\ K23\ K31 + K13\ K21\ K32 - K11\ K23\ K32 - K12\ K21\ K33 + K11K22K33$$

E di nuovo il risultato della sommatoria di questi nove determinanti è 2D.

E infine il terzo gruppo di nove determinanti

$$(19)\begin{vmatrix} K[3,3] & K[3,1] & K[3,1] \\ K[1,3] & K[1,1] & K[1,1] \\ K[1,3] & K[1,1] & k[1,1] \end{vmatrix};\ (20)\begin{vmatrix} K[3,3] & K[3,1] & K[3,2] \\ K[1,3] & K[1,1] & K[1,2] \\ K[2,3] & k[2,1] & k[2,2] \end{vmatrix};\ (21)\begin{vmatrix} K[3,3] & K[3,1] & K[3,3] \\ K[1,3] & K[1,1] & K[1,3] \\ k[3,3] & k[3,1] & k[3,3] \end{vmatrix}$$

$$(22)\begin{vmatrix} K[3,3] & k[3,2] & k[3,1] \\ k[2,3] & k[2,2] & k[2,1] \\ k[1,3] & k[1,2] & k[1,1] \end{vmatrix};\ (23)\begin{vmatrix} k[3,3] & k[3,2] & k[3,2] \\ k[2,3] & k[2,2] & k[2,2] \\ k[2,3] & k[2,2] & k[2,2] \end{vmatrix};\ (24)\begin{vmatrix} k[3,3] & k[3,2] & k[3,3] \\ k[2,3] & k[2,2] & k[2,3] \\ k[3,3] & k[3,2] & k[3,3] \end{vmatrix}$$

$$(25)\begin{vmatrix} k[3,3] & k[3,3] & k[3,1] \\ k[3,3] & k[3,3] & k[3,1] \\ k[1,3] & k[1,3] & k[1,1] \end{vmatrix};\ (26)\begin{vmatrix} k[3,3] & k[3,3] & k[3,2] \\ k[3,3] & k[3,3] & k[3,2] \\ k[2,3] & k[2,3] & k[2,2] \end{vmatrix};\ (27)\begin{vmatrix} k3,3] & k[3,3] & k[3,3] \\ k[3,3] & k[3,3] & k[3,3] \\ k[3,3] & k[3,3] & k[3,3] \end{vmatrix}$$

In questo gruppo di nove, abbiamo che

I seguenti determinanti sono nulli:

Tutti gli elementi eguali in (27)

Due colonne righe e due colonne eguali in (19), (21), (23), (24), (25), (26)

Invece due determinanti sono eguali: rispettivamente il (20) e il (22)

D= K33 K11 K22 + K31 K12 K23 + K32 K13 K21 – K33 K12 K21 – K31 K13 K22 - K32 K11 K23

E ancora una volta il risultato della sommatoria di questi nove determinanti è 2D.

Facendo la sommatoria abbiamo dunque **sei volte** il valore desiderato del determinante D, per cui, *interpretando come 3! il valore 6,* possiamo dire che il quarto elemento dello sviluppo del determinante è la sommatoria su tre indici (destinata a diventare un integrale triplo) divisa per 3!. Inoltre, il segno del termine è dato dal prodotto dei fattori ($-\lambda$) che vi compaiono, cioè da $(-\lambda)^3$.

A questo punto, identificando 1 con x, 2 con y e 3 con z, possiamo finalmente scrivere

(24):

$$D(\lambda) = 1 - \lambda \int_0^1 k(x,x,)dx + \frac{1}{2}\lambda^2 \iint_0^1 \begin{vmatrix} k(x,x) & k(x,y) \\ k(y,x) & k(y,y) \end{vmatrix} dx\, dy - \frac{\lambda^3}{3!} \iiint_0^1 \begin{vmatrix} k(x,x) & k(x,y) & k(x,z) \\ k(y,x) & k(y,y) & k(y,z) \\ k(z,x) & k(z,y) & k(z,z) \end{vmatrix} dx\, dy\, dz \;....$$

E immaginare il resto, che si ottiene costruendo intorno al determinante di ordine n, un bordo che contiene la nuova variabile.

Con questo abbiamo il denominatore D(λ) della (17). Qui, molti testi abbandonano chi si arrampica per comprendere la soluzione di Fredholm, e in particolare la formazione del determinante D(x, ξ; λ), e molti gli comunicano che la creazione è ovvia. A me non pare. Ho cercato il metodo più semplice possibile per arrivarci , e sono arrivato ad una soluzione, che può forse indicare come la matematica non sia fatta solo di numeri e formule algebriche, ma permetta di inventarsi dei formalismi nuovi, purché servano allo scopo (Il modello è preso dal bel testo ***Mathematical Methods of Physics***, di **J. Mathews e R.L. Walker**; Addison-Wesley, 1969).

II. La serie di Liouville-Neumann

Per prima cosa, però, mostriamo un modo alternativo di giungere alla soluzione della nostra equazione integrale (di Fredholm di seconda specie), cambiando il nome della variabile "muta" da ξ in y, il che è irrilevante, perché in integrazione scompare:

(25) $$\varphi(x) = f(x) + \lambda \int_a^b k(x,y)\varphi(y)dy$$

L'idea più diretta per risolvere l'equazione è di procedere per iterazioni: si incomincia col supporre l'incognita $\varphi(x) = f(x)$, e la si sostituisce *nell'integrale* in luogo di $\varphi(x)$, cioè porremo, come zeresima iterazione,

(26) $$\varphi_0(x) = f(x)$$

La prima iterazione sarà quindi:

(27) $$\varphi_1(x) = f(x) + \lambda \int_a^b k(x,y)\, \varphi_0(x)dy = f(x) + \lambda \int_a^b k(x,y)f(y)dy$$

La seconda iterazione sarà

(27 b) $$\varphi_2(x) = f(x) + \lambda \int_a^b k(x,y)\, \varphi_1(y) dy \text{ (e, sostituendo)}$$

(27 c) $$\varphi_2(x) = f(x) + \lambda \int_a^b k(x,y)\, [f(y)dy + \lambda \int_a^b k(y,y')f(y')dy']\, dy =$$

$$= f(x) + \lambda \int_a^b k(x,y)\, f(y)dy + \lambda^2 \int_a^b dy \int_a^b dy'\; k(x,y)k(y,y')f(y')dy' \ldots.$$

La continuazione è semplice: basti notare che ogni nuova interazione da *n* a *n+1* introdurrà:

- Un nuovo fattore λ, per cui si passerà da λ^n a λ^{n+1}
- Il fattore λ^{n+1} moltiplicherà un nuovo integrale contenente un numero n+1 di nuclei
- Di questi nuclei, i primi n resteranno invariati; quello aggiunto avrà per prima variabile l'ultima della serie esistente, e introdurrà una nuova variabile di integrazione.
- Se si preferisce, l'*ultimo* nucleo avrà per seconda variabile quella che era la seconda variabile nell'equazione originale (qui sarà y). Ciò non ha grande importanza, perché , dato che in seguito si integrerà anche su questa variabile, il nome non fa differenza.

Si noterà che l'unica variabile su cui non si integrerà sarà la x di partenza, come deve essere, se vogliamo avere come risultato una funzione di x, e non una costante.

Il Nucleo $k_m(x,y)$ risulta essere $\ldots \int_a^b k(x,\xi_1)k(\xi_1,\xi_2) \ldots k(\xi_{m-1},y)f(y)\, d\xi_1\; d\xi_2 \ldots dy$

E sinteticamente, si potrà scrivere:

(28) $$\varphi(x) = f(x) + \lambda \int_a^b k(x,\xi)\, f(\xi)d\xi + \sum_{m=2}^{\infty} \lambda^m \iiint \ldots \int_a^b k(x,\xi_1)k(\xi_1,\xi_2) \ldots k(\xi_{m-1},\xi_m)f(y)\, d\xi_1\; d\xi_2 \ldots dy$$

Il metodo è dovuto a **Liouville (1837) e Neumann (1870)**, e potremo scrivere la soluzione come:

(29) $$\varphi(x) = f(x) + \lambda \int_a^b H(x,y;\, \lambda) f(y)dy$$

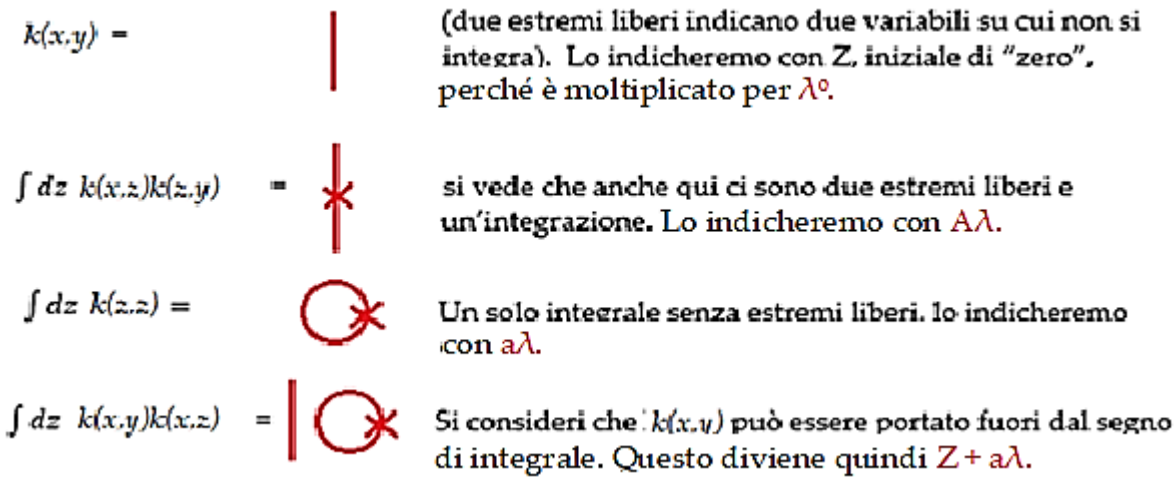
Qui, il **nucleo "risolvente"** in cui si può condensare l'intera serie da λ a λ^m, sarà posto uguale a $H(x,y;\, \lambda)$. La serie non converge sempre: occorre che λ sia sufficientemente piccolo, e il nucleo K limitato (cioè non possa assumere valori infinti nel dominio di integrazione). *Ma qui non ci addentreremo in questioni di convergenza.* Pur presentando, come ogni problema matematico, un interesse intrinseco, i teoremi che dimostrano la convergenza delle serie erano particolarmente importanti quando un matematico, che disponeva di mezzi di calcolo limitati, doveva addentrarsi nel calcolo di una serie, senza sapere se questa si sarebbe dimostrata convergente o no: esistono ora mezzi di calcolo che

permettono di verificare direttamente e rapidamente se la serie converge o minaccia di divergere, o diverge. Allora, invece, era più utile servirsi di *criteri di convergenza* applicabili se possibile in quatttro e quattr'otto.

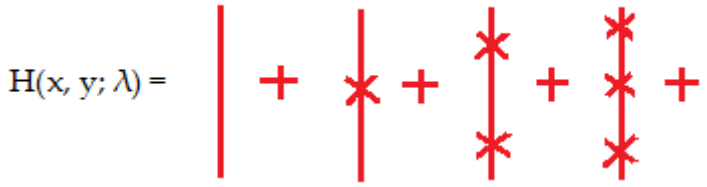
Notiamo che il vantaggio di aver isolato nella (17) il termine f(x) fuori dell'integrale renderà possibile paragonare in modo utile i due nuclei risolventi.

III: Un nuovo formalismo e i determinanti di Fredholm.

Ora si può introdurre la prima parte del formalismo a cui accennavo per trovare il numeratore del nucleo risolvente di Fredholm. Si considerino i seguenti diagrammi:

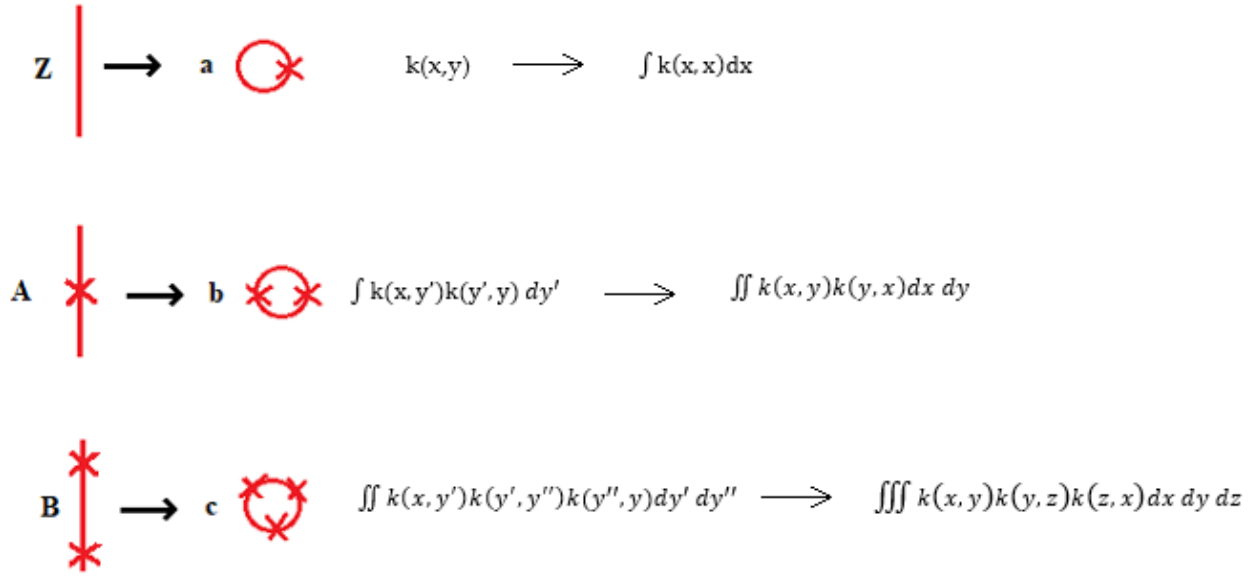


Da tutto ciò, per incominciare, si deduce che la ***serie di Liouville-Neumann per il nucleo risolvente*** (non per la *soluzione)* può essere scritta nel seguente modo (30), che lascia sempre due estremi liberi, che corrispondono alle variabili x e y. La *soluzione* (29) sarà ottenuta eseguendo l'integrazione su uno di essi (cioè l'estremo y, se sarà stato lasciato per ultimo), previo prodotto per f(y):



Con ovvia notazione possiamo scrivere: $H(x,y;\lambda) = Z + (A\lambda) + (B\lambda^2) + (C\lambda^3) \ldots$ (dove le maiuscole indicano che il diagramma è composto di segmenti rettilinei, in cui le lettere, prese in ordine alfabetico, indicano il numero di crocette, ciascuna delle quali corrisponde a ***un'integrazione e a un fattore λ***. In particolare, Z è l'iniziale del numero zero di crocette e integrazioni (e di fattori $+\lambda$).

Aggiungiamo una seconda notazione: “chiudendo in cerchio” un termine maiuscolo, si “annoda” aggiungendo una crocetta e si ottiene un termine minuscolo, indicato con un cerchio, dell’ordine superiore, perché ogni “nodo” aggiunge una crocetta:



Il denominatore $D(\lambda)$

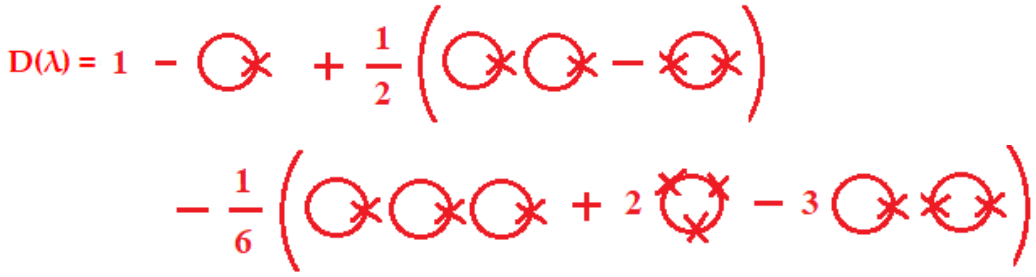
Applichiamo ora la notazione al denominatore (25)

$$D(\lambda) = 1 - \lambda\int_0^1 k(x,x,)dx + \frac{1}{2}\lambda^2 \iint_0^1 \begin{vmatrix} k(x,x) & k(x,y) \\ k(y,x) & k(y,y) \end{vmatrix} dx\,dy - \frac{1}{3!}\iiint_0^1 \begin{vmatrix} k(x,x) & k(x,y) & k(x,z) \\ k(y,x) & k(y,y) & k(y,z) \\ k(z,x) & k(z,y) & k(z,z) \end{vmatrix} dx\,dy\,dz \;\ldots$$

i cui primi tre termini, svolgendo il determinante di secondo ordine, divengono

(25 b) $D(\lambda) = 1 - \lambda\int_0^1 k(x,x,)dx + \frac{1}{2}\lambda^2 \iint_0^1 \big(k(x,x)k(y,y) - k(x,y)k(y,x)\big)dx\,dy - \cdots$

ricordando che un’integrazione è indicata da una croce, che si porta dietro un fattore (-λ), possiamo scrivere (o disegnare?), in termini del numero di crocette , che sono le potenze di λ, la seguente serie:



che, con ovvia notazione, può essere scritta

(25c) $D(\lambda) = 1 - a\lambda + (\tfrac{1}{2})((a\lambda)(a\lambda) - b\lambda^2) - (1/6)((a\lambda)(a\lambda)(a\lambda) + 2c\lambda^3 - 3(a\lambda)b\lambda^2))\ldots$

Il terzo termine non dovrebbe presentare problemi: si svolga il determinante, per esempio seguendo la ben regola di **Sarrus** (1), valida per i determinanti 3 x 3 e non oltre, e si identifichino i temini simili.

Può valer la pena ricordare che il termine *a* λ, viene ottenuto "annodando" gli estremi di Z con una crocetta, che introduce un fattore $-\lambda$. In effetti si ha

$$k(x,y) \rightarrow -\lambda \int k(x,x)dx$$

Questa semplice operazione riassume tre azioni: l'identificazione degli estremi della Z, e l'integrazione, moltiplicata per un fattore -λ.
Invece, dal termine A λ , annodando gli estremi, cioè dando lo stesso nome alle due variabili e integrando, il che aggiunge un fattore λ , si ottiene b $(-\lambda)^{2}$, e via dicendo.

$$\iint k(x,z)k(z,x)dx\,dz = \iint k(z,x)k(x,z)dx\,dz$$

Non c'è motivo di pensare che la formula ci debba tradire per termini di ordine più alto, e lo si può verificare svolgendo i determinanti, affascinante compito che lasciamo a coloro che sono portati per questo genere di cose. Gli altri, possono credermi. I conti li ho fatti.
Il terzo termine, ad esempio, sappiamo che è un sesto di un dato determinante che ha sei termini. Usando i nostri diagrammi e tenendo conto della possibilità di variare i nomi degli indici e l'ordine nei prodotti, esso equivale a:

$$-\frac{\lambda^3}{6}\iiint dx\,dy\,dz\,[k(x,x)k(y,y)k(z,z) + 2\,k(x,y)k(y,z)k(z,x) - 3\,k(x,x)\,k(y,z)k(z,y)]$$

Si osservi che il denominatore è costituito unicamente da circoli chiusi: ciò significa che esso dipende solo da λ, poiché si è eseguita l'integrazione sulle variabili restanti.

Il numeratore $\boldsymbol{D(x,\xi;\lambda)}$.

Passiamo ora al calcolo di $D(x,\xi;\lambda)$. Dato che deve essere

(30) $$H(x,y;\ \lambda) = \frac{D(x,\xi;\lambda)}{D(\lambda)}$$

Deve anche essere:

(31) $$D(x,\xi;\lambda) = H(x,y;\ \lambda)D(\lambda)$$

In cui il membro di destra è noto.

A scopo di esempio, accontentiamoci del prodotto fino alle potenze di λ di ordine 2.
Il prodotto, con la nostra notazione, è

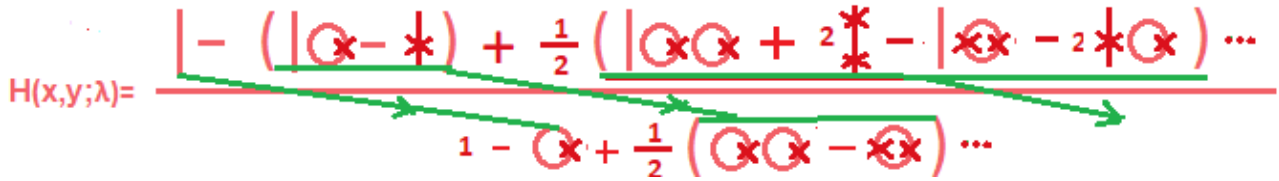
(32) $(Z + A\lambda + B\lambda^2)(1 - a\lambda + (1/2)(a\lambda a\lambda - b\lambda^2)$ ....= $\boldsymbol{Z - a\lambda Z + A\lambda - aA\lambda^2 + B\lambda^2 + \frac{1}{2}a^2\lambda^2 Z - b\lambda^2 Z} + \frac{1}{2}a^2 A\lambda^3 - Ab\lambda^3 - aB\lambda^3 + \frac{1}{2}a^2 B\lambda^4 - bBx\lambda^4$...

In tutto abbiamo 7 termini (rossi) di grado inferiore al terzo, i quali possono essere scritti, secondo le nostre regole, in ordine delle potenze di λ:

(33) $D(x,\xi;\lambda)_2 = k(x,\xi)$ - $\lambda \int dz\,(k(x,\xi)k(z,z) - k(x,z)k(z,\xi)) +$
$\frac{\lambda^2}{2}\iint dz\,dz'[k(x,\xi)\,k(z,z)k(z',z') + 2\,k(x,z)k(z,z')k(z',\xi) - k(x,\xi)k(z,z')k(z',z) -$
$2\,k(x,z)k(z,\xi)k(z',z')]$ ...

sempre tenendo conto della possibilità di variare i nomi delle variabili su cui si integra e l'ordine nei prodotti. Inoltre, si noti che K(x, ξ) esce sempre dall'integrale per D(x, ξ; λ). L'integrale su ξ verrà poi eseguito una volta che il nucleo risolvente sarà stato moltiplicato per f(ξ).

L'ispezione dei diagrammi e delle regole di formazione dei medesimi, ci suggerisce la regola che ogni termine del denominatore , eccetto la cifra 1, è ottenuto "annodando" con una X ogni elemento del precedente termine del numeratore, dividendo per il numero di nuclei K (le crocette) presenti in ogni addendo, e cambiando segno. Quanto all'esponente di (-λ), esso sarà eguale al numero delle crocette.

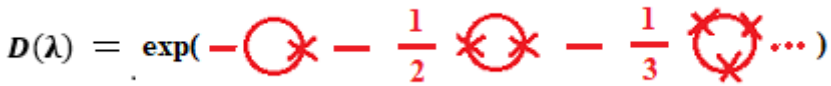


La creazione del termine di terzo ordine del denominatore è suggerita come esercizio. Il risultato è in Nota 1.

Due giochi.

I. La regola data per passare da un termine del numeratore a quello del denominatore, può essere invertita, e permette di passare da un termine del denominatore (li conosciamo in linea di principio tutti) al termine corrispondente del numeratore "disannodando" ogni termine annodato, moltiplicando per il numero di crocette e cambiando segno. Questo lo si può fare senza passare per la serie di Liouville Neumann, e lo si lascia come gioco per il lettore. Si ricordi, come aiuto, che:

II: Coloro che conoscono lo sviluppo in serie dell'esponenziale exp(x) possono divertirsi a provare il notevole sviluppo



4. Per completezza.

Il mio scopo era quello di indicare in modo euristico come Fredholm trovò la soluzione della sua equazione. Tuttavia, forse i più importanti risultati che Fredholm ottenne furono tre teoremi, che non enuncerò troppo rigorosamente e dei quali non darò la dimostrazione (che di rado è data nei testi di matematica per i fisici) ma che appariranno ragionevoli a chi abbia qualche conoscenza un po' più che elementare degli autovalori e autofunzioni dei sistemi di equazioni algebriche lineari. Non è questo il luogo per trattare in maniera approfondita il problema per il caso delle equazioni algebriche. I concetti principali sono illustrati da esempi "minimali"nel complemento A.

Teoremi di Fredholm

Fredholm ricavò, non senza sforzo, tre teoremi che, detti di schianto all'inesperto, sembrano assai astrusi. Tuttavia, una volta letto il complemento A , essi possono anche apparire abbastanza prevedibili.

1. Alternativa: O l'equazione inomogenea

(14) $$\varphi(x) = f(x) + \lambda \int_a^b K(x,\xi)\varphi(\xi)d\xi$$

ha un'unica soluzione per qualsiasi funzione f(x) (il che avviene se λ non è un autovalore), oppure l'equazione omogenea:

(14c) $$\varphi(x) = \lambda \int_a^b K(x,\xi)\varphi(\xi)d\xi$$

ha almeno una soluzione non banale ($\varphi(x) = 0$) . Allora λ è un autovalore, e la soluzione è un'autofunzione.

*Sappiamo che (**A-I**) la condizione perché un sistema di equazioni algebriche lineari di primo grado abbia una soluzione unica è che il determinante dei coefficienti sia **diverso** da zero, ciò che avviene nell'equazione (12c) se **λ non è un autovalore** della matrice dei coefficienti.*

2. Se λ non è un autovalore (primo caso), allora λ non è neppure un autovalore dell'equazione trasposta:

(14) $$\varphi(x) = f(x) + \lambda \int_a^b K(\xi,x)\varphi(\xi)d\xi$$

Mentre, se λ è un autovalore (secondo caso), λ è anche un autovalore dell'equazione omogenea trasposta, cioè l'equazione

(14c) $$\varphi(x) = \lambda \int_a^b K(\xi, x)\varphi(\xi)d\xi$$

ha almeno una soluzione non banale.

Sappiamo che ***(A- II)*** *gli autovalori della matrice (quadrata) dei coefficienti sono gli stessi della matrice trasposta, cioè della matrice in cui sono state scambiate le righe colle colonne, ciò che qui corrisponde allo scambio delle due variabili. Gli autovalori sono gli stessi, ma anche qui le autofunzioni non lo sono in genere.*

3. Se λ è un autovalore, allora l'equazione inomogenea

(14) $$\varphi(x) = f(x) + \lambda \int_a^b K(x, \xi)\varphi(\xi)d\xi$$

ha una soluzione se e solo se

$$\int_a^b f(\xi)\psi(\xi)d\xi = 0$$

per ogni funzione $\psi(\xi)$ che soddisfa l'equazione omogenea trasposta.

(14c) $$\varphi(x) = \lambda \int_a^b K(\xi, x)\varphi(\xi)d\xi$$

E noi sappiamo che (A-**III)** *Se il determinante dei coefficienti è nullo, una soluzione del sistema non omogeneo è possibile purché il termine noto sia un vettore ortogonale all'autovettore della matrice trasposta corrispondente allo stesso moltiplicatore.*

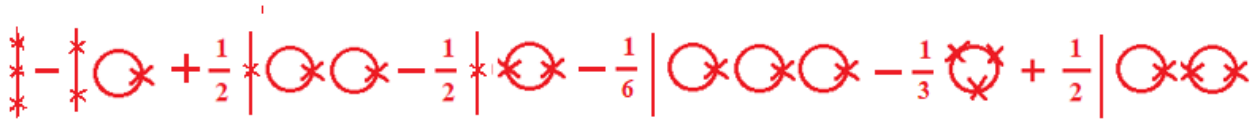
Vediamo quindi che i tre teoremi di Fredholm possono essere considerati come estensioni al continuo dei teoremi del calcolo matriciale elementare.

5. Commiato.

E con questo ho terminato il mio lavoro divulgativo sulle equazioni di Fredholm. Spero che sia stato utile a qualcuno.

NOTE:

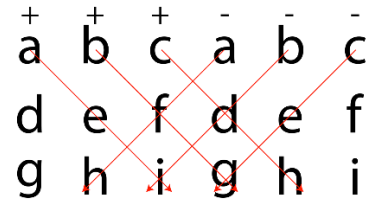
(1) Fattori di ($-\lambda^3$) – Esercizio a pag.30
Al numeratore:



Lo si può ottenere sia col metodo simbolico, sia coll'algoritmo (32), continuati fino a includere tutti i grafici con tre crocette, o tutte le potenze ($-\lambda^3$).

Al denominatore: il risultato è a pag. 45.

Regola di Sarrus (da Wikipedia):



aei + bfg + cdh - afh - bdi - ceg