Capitolo 4

Presentazione e analisi grafica dei dati

L'esecuzione di un esperimento consiste, in molti casi, nel variare una grandezza, il cui valore è deciso dallo sperimentatore, e nell'osservare le variazioni delle grandezze fisiche che ci si aspetta possano dipendere da quella variata. Come esempio, consideriamo la misura della costante elastica di una molla; l'esperimento consiste nella variazione della massa che si applica ad un estremo della molla e nella misura del conseguente allungamento; la grandezza che lo sperimentatore varia è la massa applicata alla molla e la grandezza si misura è l'allungamento della molla. Per una analisi preliminare del fenomeno è opportuno creare delle tabelle in cui si riportano in modo quanto più possibile ordinato, i valori delle grandezze misurate. Per una compilazione ordinata di una tabella si intende che i valori delle grandezze variate dallo sperimentatore siano riportate in modo monotòno (crescente o decrescente), annotando unità di misura e incertezze delle singole misure. Una lettura attenta delle tabelle spesso mette in evidenza banali errori di trascrizione o incompletezze nell'esecuzione dell'esperimento. La tabella 4.1 è un esempio concreto su come si possano organizzare i dati acquisiti nell'esperimento di allungamento della molla nel quale si studia sperimentalmente l'andamento dell'allungamento di una molla in funzione del peso applicato.

Tabella 4.1: Dati dell'esperimento sull'allungamento di una molla. Nella tabella sono riportate nelle due colonne di sinistra le misure della grandezza variata dello sperimentatore, cioè la massa applicata alla molla con la sua incertezza e nelle altre due colonne la misura della grandezza "dipendente" (l'allungamento della molla) con la sua incertezza. Si noti che il numero di cifre significative di ogni misura è coerente con il valore della sua incertezza .

Massa		Coordinata	
m (g)	$u_m(g)$	x (cm)	u_x (cm)
10.0	0.1	22.7	0.4
20.0	0.1	25.1	0.4
30.0	0.1	27.9	0.4
40.0	0.1	30.7	0.4
50.0	0.1	33.6	0.4
60.0	0.1	36.9	0.4

4.1 Rappresentazione grafica dei dati

Il primo utilizzo dei dati raccolti in una tabella è quello di costruire un grafico. I grafici sono una importante modalità per la presentazione e l'analisi preliminare dei dati scientifici. I grafici possono essere usati per suggerire relazioni fisiche, per confrontare i dati sperimentali con un andamento matematico previsto e per determinare facilmente alcuni parametri quali la pendenza di una linea retta. Infatti l'occhio umano riesce a distinguere molto efficacemente se dei punti sono allineati tra loro. Descriviamo nel seguito i vari tipi di grafico che si utilizzano in fisica.

4.1.1 Grafici lineari

Nel grafico lineare entrambi gli assi, ascisse e ordinate, hanno scale lineari. Per la creazione di un grafico lineare[1] è opportuno seguire una specifica sequenza di passi che sono validi sia se si riportano manualmente i dati su un foglio reale di carta millimetrata sia se, come sta diventando prassi, si utilizzano programmi di grafica computerizzata per l'esecuzione dei grafici[2]. Le operazioni da eseguire per la compilazione dei grafici, che la pratica renderà automatiche, sono elencate qui di seguito:

1. Organizzare in una tabella i dati da inserire nel grafico (come già detto precedentemente) .

2. Decidere quale grandezza è da mettere sull'asse x (ascisse); di solito nell'asse x si pone la variabile indipendente e in quello y (ordinate) la variabile dipendente.

3. Decidere se l'origine deve apparire sul grafico. In alcuni casi è richiesto che l'origine appaia nel grafico, anche se il valore "zero" non è parte dei dati; per esempio, nel caso in cui si voglia determinare graficamente un'intercetta.

4. Scegliere una scala per ogni asse, ovvero determinare il valore iniziale da da cui parte la scale e a quanti centimetri corrisondono un numero opportuno di unità di misura della grandezza da riportare nel grafico (Esempio: valore iniziale 0 grammi e a 5 grammi corrispondono 20 cm). Le scale devono essere scelte in modo che i dati da graficare si estendano su quasi tutta la carta millimetrata, e inoltre la scelta della scala deve essere "razionale" ovvero in modo da rendere facile localizzare quantità arbitrarie sul grafico. (Esempio: 5 divisioni = 23 centimetri è una pessima scelta.) Solo le divisioni principali devono essere riportate su ciascun asse.

5. Scrivere accanto a ciascun asse la variabile che indica la grandezza rappresentata *con le sue unità*. E' buona pratica dare un titolo al grafico, scrivendolo nel margine superiore del grafico eventualmente aggiungendo la data nella quale sono stati raccolti i dati. (Esempio: "Allungamento della molla vs massa, Data: gg/mmm/yyyy")

6. Marcare ogni punto sperimentale nel grafico. Lo stile consigliato è marcare un punto circondato da un piccolo cerchio ($\odot$). Equivalentemente si può usare il segno per ($\times$) o il segno più ($+$).

[1] La sequenza di azioni descritte è valida oltre che per i grafici lineari anche per gli altri tipi di grafici che verranno descritti nei prossimi paragrafi.

[2] Usando programmi di *computer* alcuni passi, come ad esempio la scelta delle scale degli assi sono fatte in modo automatico, anche se non sempre nel modo migliore. Tuttavia altre azioni, come il riportare nome ed unità delle variabili sugli assi, richiedono sempre un intervento umano.

7. Non unire i punti sperimentali con una linea spezzata. La "spezzata" rende più difficile capire l'andamento generale della relazione fra le due grandezze e inseguendo le fluttuazioni casuali dei dati può indurre a ipotizzare andamenti inesistenti.

8. Nel caso molto comune in cui si ipotizzi un andamento lineare fra le grandezze misurate si può utilizzare un righello trasparente per tracciare ad "occhio" una ragionevole retta in modo da lasciare un numero uguale di punti su ciascun lato della linea. Questo modo di procedere per quanto non rigoroso porta spesso ad un risultato molto vicino a quello che si ottiene con più complesse procedure di elaborazione matematica dei dati che saranno spiegate nel seguito.

9. Per determinare la pendenza della linea, scegliere due punti sulla linea i cui valori sono facilmente leggibili e che si estendano per quasi l'intera larghezza del grafico. Questi punti non dovrebbero essere, se non per caso, i punti dei dati sperimentali. Si ricordi che la pendenza ha unità di misura pari al rapporto tra le unità di misura dei due assi.

10. L'incertezza sul valore della pendenza può essere stimata come rapporto tra la maggiore delle incertezze dei due punti finali, e la distanza tra i due punti stessi.

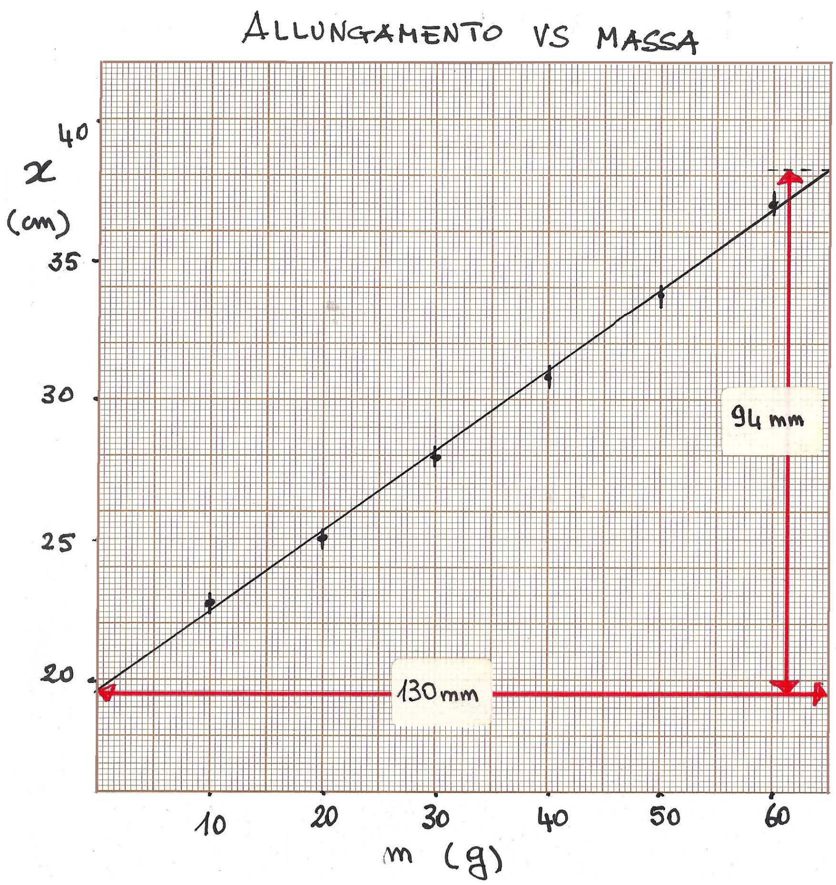


Figura 4.1: Grafico con scale lineari dei dati di tabella 4.1. Nel grafico sono indicate le misurazioni effettuate con un righello millimetrato e quindi rapportate alle scale scelte per la valutazione del coefficiente angolare della retta tracciata. Vedi il testo per ulteriori dettagli.

Come applicazione delle regole esposte creiamo un grafico in carta lineare con i dati della tabella 4.1. Nell'asse delle ascisse poniamo il valore della massa applicata alla molla scegliendo una scala in cui 10 mm corrispondono a 5 g e nell'asse delle ordinate riportiamo l'allungamento scegliendo una scala in cui 10 mm corrispondono a 0.2 cm di allungamento della molla. Nella figura 4.1 sono riportati i punti sperimentali ed è stata tracciata "ad occhio" una retta che approssima i dati. La teoria prevede infatti che per una molla in equilibrio ci sia una relazione lineare tra la forza applicata F (nel nostro caso $F = mg$) e il conseguente allungamento $x - x_o$: $F = k(x - x_o)$ con k *costante della molla* e x_o lunghezza della molla a riposo (quando $F = 0$). La retta disegnata ha equazione $x = (g/k)m + x_o$ e quindi, utilizzando la retta tracciata nel grafico, la stima del parametro k è:

$$k = \frac{\Delta F}{\Delta x} = g\frac{\Delta m}{\Delta x} = g\frac{80\text{mm}\, C_m}{44\text{mm}\, C_x}$$

dove C_m e C_x sono fattori di conversione; il primo trasforma i millimetri misurati nell'asse delle ascisse in kg il secondo trasforma i millimetri misurati nell'asse delle ordinate nell'allungamento espresso in metri. Dai dati sulle le scale delle ascisse e delle ordinate è facile calcolare i valori di C_m e di C_x:

$$C_m = 1\frac{\text{g}}{\text{mm}}10^3\frac{\text{kg}}{\text{g}} = 10^3\frac{\text{kg}}{\text{mm}}, \qquad C_x = 5\frac{\text{cm}}{\text{mm}}10^2\frac{\text{m}}{\text{cm}} = 500\frac{\text{m}}{\text{mm}}$$

Sostituendo:

$$k = g\frac{80\text{mm}\, C_m}{44\text{mm}\, C_x} = 9.8 \times \frac{80 \times 10^3}{44 \times 500} = 35.6 \text{ kg s}^{-2}$$

Si lascia al lettore come esercizio la stima del valore di x_o.

Linearizzazione. Come mostrato anche nel precedente esempio l'uso dei grafici evidenzia facilmente una dipendenza lineare tra le grandezze fisiche riportate sugli assi per cui, come già notato nel punto 7 del paragrafo precedente, alle volte è opportuno *linearizzare* la relazione matematica che descrive i dati. Cosicché se dovessimo riportare su di un grafico una relazione matematica tra spazio e tempo in un moto uniformemente accelerato, descritta dall'equazione $s = (1/2)at^2$, potremo riportare nell'asse delle ascisse i valori di t^2 e su quelle delle ordinate i valori di s. In questo modo si ottiene una dipendenza lineare tra le variabili s e t^2 il cui coefficiente angolare è $a/2$. Operando in questo modo diremo che abbiamo linearizzato la relazione studiata.

4.1.2 Grafici Logaritmici

Oltre alle scale lineari sono molto utilizzate le scale logaritmiche che hanno vari scopi. Uno è quello di linearizzare certe classi di funzioni molto utilizzate nella descrizione di fenomeni fisici, l'altro è quello di riportare in uno stesso grafico diversi ordini di grandezza di una variabile.

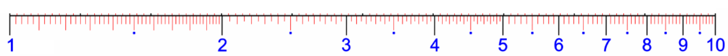


Figura 4.2: Scala logaritmica

Una scala logaritmica è un asse sul quale sono riportati segmenti di lunghezza proporzionale al logaritmo dei numeri rappresentati. Di solito gli interi sono indicati sulla scala come mostrato nella figura 4.2. Fissata la base a del logaritmo, al numero 1 corrisponde un segmento di lunghezza nulla, ovvero l'origine dell'asse logaritmico; al numero 2 corrisponde un segmento di lunghezza $\log_a 2$, al numero 10 corrisponde un segmento $\log_a 10$, al numero 100 corrisponde un segmento di lunghezza $\log_a 100 = 2\log_a 10$, al numero 20 corrisponde un segmento di lunghezza $\log_a 20 = \log_a 2 + \log_a 10$, e così via. Si noti che, poiché $\log_a 10^n = n\log_a 10$, la lunghezza del segmento che rappresenta una "decade", da 1 a 10 , da 10 a 100, da 100 a 1000,. . . ha sempre la stessa lunghezza.

La scelta della base del logaritmo corrisponde ad un semplice fattore di proporzionalità della scala, infatti, utilizzando le proprietà dei logaritmi:

$$\log_a x = \frac{\log_b x}{\log_b a} = \log_a b \log_b x = \text{cost } \log_b x$$

Le scale logaritmiche sono usate nei diagrammi bidimensionali del tipo detto semi-logaritmico(un

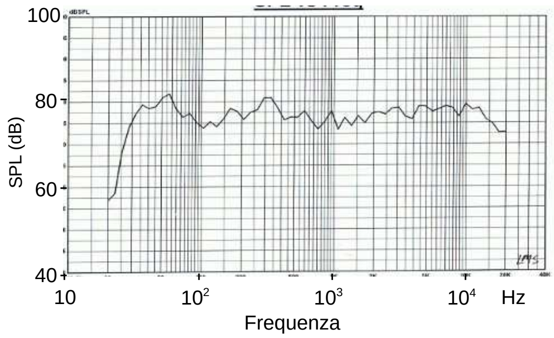


Figura 4.3: Esempio di grafico semi-log che mostra la SPL, misura dell'intensità sonora di un altoparlante, in funzione della frequenza nell'intervallo 10 Hz – 40 kHz. Si noti che l'uso della scala logaritmica permette di rappresentare nell'asse delle ascisse tre ordini di grandezza.

asse lineare e un asse logaritmico) oppure del tipo detto doppio-logaritmico (entrambi gli assi logaritmici). Esistono delle carte millimetrate prestampate con i detti diagrammi anche se oggi il loro utilizzo è stato soppiantato dalla grafica computerizzata.

Grafico semi-logaritmico. In un grafico semi-logaritmico (alle volte abbreviato in semi-log) l'asse con scala logaritmica può essere quello delle ascisse o quello delle ordinate; è nel primo caso il grafico è tipicamente utilizzato per comprimere la scala in modo da contenere vari ordini di grandezza come ad esempio è mostrato nella figura 4.3 .

Il grafico semi-log con l'asse delle ordinate ha, oltre alla compressione della scala (vedi ad esempio nella figura 4.4), la proprietà di linearizzare gli andamenti esponenziali del tipo $y = C\mathrm{e}^{kx}$ dove C e k sono delle costanti[3]. Infatti prendendo il logaritmo di entrambi i membri di questa equazione si ha

$$\log_a y = \log_a C\mathrm{e}^{kx} = kx\log_a \mathrm{e} + \log_a C \tag{4.1}$$

Come si nota da questa equazione il logaritmo di y è proporzionale a x e quindi il grafico della relazione (4.1) avrà un andamento rettilineo se tracciato in una carta millimetrata semilogaritmica (vedi la figura 4.5). Per valutare k nella (4.1) si può procedere in due modi: per via algebrica o per via grafica.

[3]Si noti che se x è una grandezza fisica con le sue dimensioni, la costante k deve avere le dimensioni inverse per quanto detto nel paragrafo 1.8.2

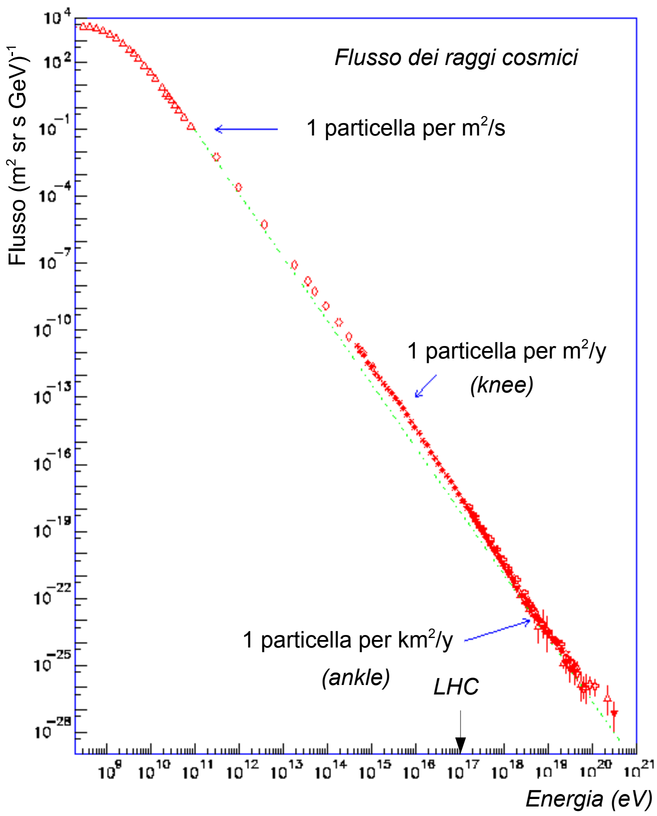


Figura 4.4: Questo grafico illustra come la scala logaritmica metta in evidenza strutture dei dati che l'uso di una scala lineare non sarebbe in grado di mostrare. Questo grafico, noto come *"All particle spectrum*, rappresenta il numero dei raggi cosmici che colpiscono l'atmosfera terrestre per unità di superficie, di angolo solido, di tempo e di energia in funzione dell'energia.

Via algebrica. Si scelgono due punti , (x_1, y_1) e (x_2, y_2) sulla retta quanto più possible lontani tra loro per minimizzare le incertezze relative di di lettura e si ricava il parametro k come:

$$k = \frac{\ln y_2 - \ln y_1}{x_2 - x_1}$$

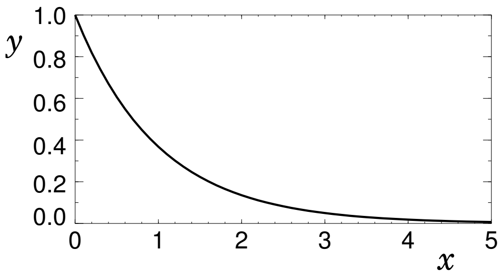
avendo scelto come base dei logaritmi quella naturale.

Via grafica. Per via grafica, si deve misurare con un righello la distanza Δy in millimetri (vedi la figura 4.5) e rapportala all'unità di misura scelta (tipicamente si sceglie la lunghezza in millimetri di una decade L_{Decade}). Questa misura è la differenza dei logaritmi nella base scelta. In formule:

$$k = \frac{\Delta y}{L_{Decade}} \cdot \frac{1}{x_2 - x_1} \ln 10$$

Grafico doppio-logaritmico. I grafici nei quali i due assi, delle ascisse e delle ordinate, hanno scale logaritmiche si chiamano doppio-logaritmici (abbreviato doppio-log). In questi grafici una funzione di potenza appare come una retta. Consideriamo una relazione tra grandezze del tipo

$$y = Ax^\alpha \tag{4.2}$$



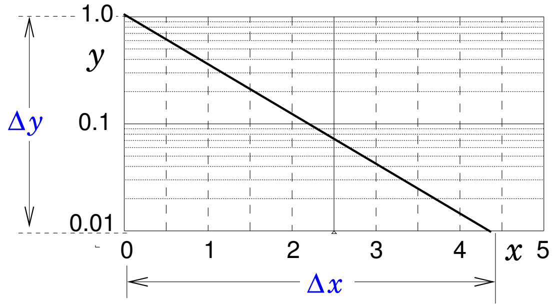


Figura 4.5: Grafico della funzione $y = \mathrm{e}^{-kx}$ (con $k = 1$) in scala lineare (a sinistra) e in scala semi-logaritmica (a destra). Nel grafico di destra è mostrato come si possa stimare graficamente la costante k che appare nella definizione della funzione.

Tabella 4.2: Periodo di rivoluzione in anni (y) e distanza media dal sole in AU (Astronomic Units) dei pianeti del sistema solare (dati NASA).

Pianeta	Periodo T (y)	R (AU)
Mercurio	0.2411	0.387
Venere	0.6156	0.723
Terra	1.001	1.00
Marte	1.882	1.52
Giove	11.87	5.21
Saturno	29.44	9.58
Urano	83.81	19.20
Nettuno	163.8	30.05
Plutone	248.1	39.48

prendendo i logaritmi di entrambi i membri si ha

$$\log y = \alpha \log x + \log A \tag{4.3}$$

Questa relazione rappresenta una retta nelle variabili $\log y$ e $\log x$ con coefficiente angolare α e intercetta $\log A$. Si noti che il coefficiente angolare di questa retta è l'esponente della variabile x nella (4.2). Per la valutazione dei parametri nella (4.3) si usano le stesse procedure indicate nel paragrafo precedente per i grafici semi-logaritmici. Come esempio dell'uso di un grafico doppio-logaritmico per verificare la veridicità di una legge fisica prendiamo in considerazione la terza legge di Keplero applicata ai dati del nostro sistema solare planetario. La terza legge di Keplero dice che il quadrato del periodo orbitale di un pianeta T è proporzionale al cubo della sua distanza media dal Sole R. In formule

$$T^2 \propto R^3 \qquad \text{oppure, isolando } R: \quad R \propto T^{2/3} = T^{0.66} \tag{4.4}$$

La relazione (4.4) è una legge di potenza del tipo (4.2) per cui in un grafico doppio-log i dati di periodo e distanza media dal sole dei pianeti si dovrebbero allineare per conferemare la terza legge di Keplero. In tabella 4.2 sono riportati i dati planetari presi da sito della NASA.

L'ispezione dei dati in tabella indica che nell'asse delle ascisse, dove metteremo il periodo T ci devono essere almeno 4 decadi e nell'asse delle ordinate almeno tre decadi se desideriamo

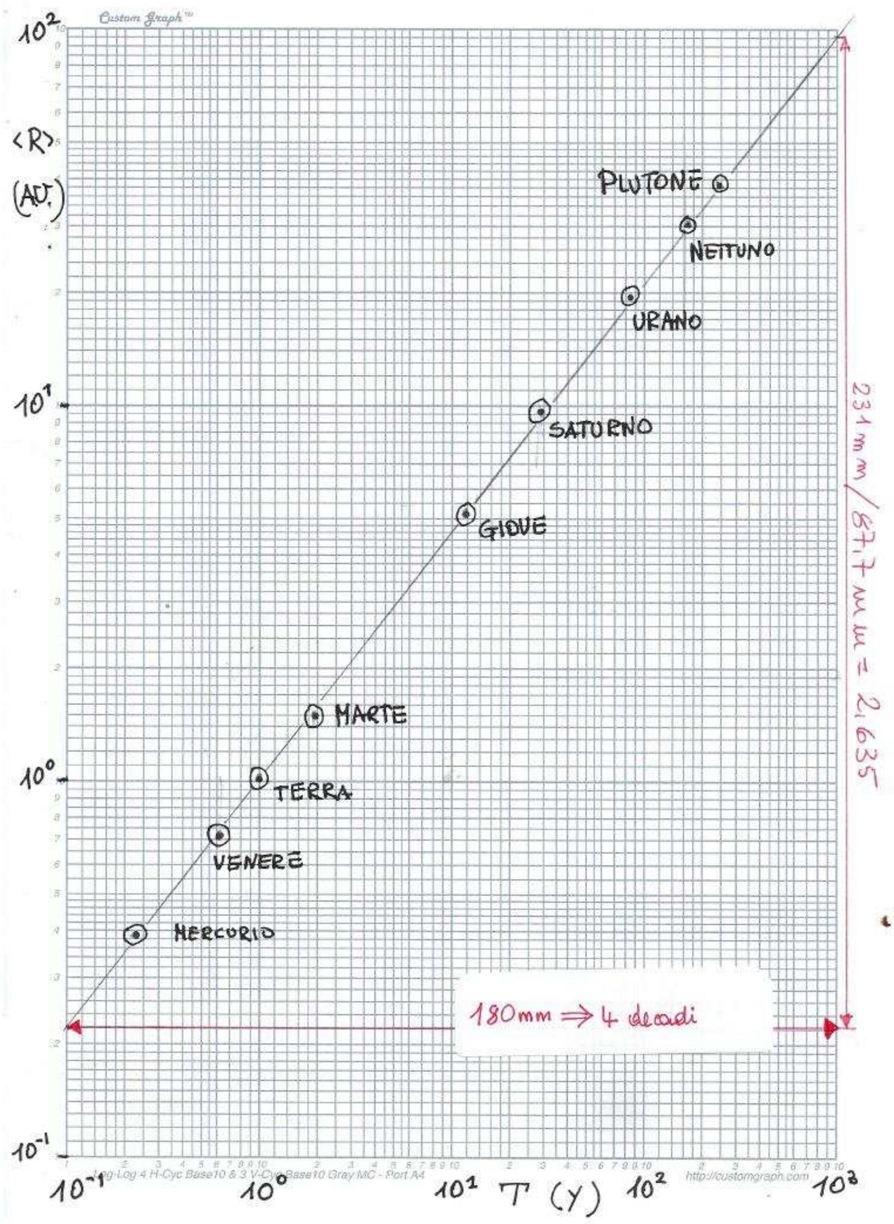


Figura 4.6: Grafico doppio-log che mostra la verifica della terza legge di Keplero con i dati del sistema solare (vedi testo).

che tutti i punti sperimentali stiano nel grafico. Fissato il numero di decadi per il grafico doppio-log[4] con il criterio appena esposto, riportiamo i punti sperimentali su grafico e quindi si traccia una retta cercando di farla passare per tutti i punti sperimentali come mostrato nella figura 4.6. La retta tracciata deve essere prolungata attraverso tutto il foglio per poter misurare nel modo migliore la sua inclinazione. La misurazione dell'inclinazione della retta disegnata nel grafico è stato ottenuta per via grafica (vedi il paragrafo 4.1.2) con il seguente

[4]E' possibile scaricare dalla rete carte millimetrate lineari, logaritmiche e doppio logaritmiche scegliendo il numero di decadi desiderate nei due assi

risultato:

$$\alpha = \frac{\Delta y}{\Delta x} = \frac{2.635 \text{ decadi}}{4.0 \text{ decadi}} = 0.66$$

valore in ottimo accordo con qiello previsto la terza legge di Keplero.

4.1.3 Altri tipi di grafico – Grafici polari

Le funzioni che dipendono da un angolo piano possono essere rappresentate anche in modo polare. Questa modalità si basa sul sistema di coordinate polari del piano in cui ogni punto P del piano è identificato dalla sua distanza r da un punto O, detto polo, e dall'angolo θ tra il segmento $\overline{OP}$ e una prefissata semiretta uscente dal polo O, asse del sistema. Il grafico polare della funzione $r = f(\theta)$ si ottiene tracciando sulla semiretta uscente dal polo con un angolo θ rispetto all'asse, un punto ad una distanza dal polo proporzionale a r. Come esempio di un grafico polare consideriamo la gittata R di un proiettile che parta con velocità in modulo pari a v_o in funzione di θ, angolo di tiro o "alzo". E' un facile esercizio di meccanica arrivare alla relazione $R = 2v_o^2 \sin 2\theta/g$. Il grafico polare di R in funzione di θ è mostrato nella figura 4.7.

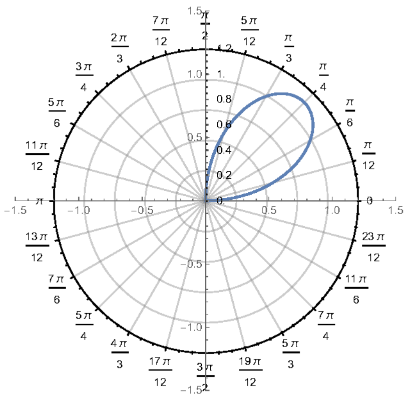


Figura 4.7: Esempio di grafico polare. Il grafico mostra l'andamento della gittata di un proiettile che parte con velocità in modulo costante in funzione dell'angolo di "alzo" , limitato nell'intervallo $[0, \pi/2]$. Per questo grafico si è scelto $v_o = 7.5\,\text{ms}^{-1}$ Come è noto e come si deduce dal grafico, la massima gittata si ottiene per $\theta = \pi/4$.

4.1.4 Istogrammi

Le rappresentazioni grafiche illustrate nei paragrafi precedenti sono utilizzate nello studio di dipendenze funzionali di una grandezza in funzione di un'altra come ad esempio è rappresentato nelle figure 4.1, 4.5 e 4.6.

Nelle misurazioni ripetute di una grandezza fisica continua, come quella citata nell'esempio 1 del paragrafo 3.7 oppure quella mostrata nella figura 3.1, l'intento non è quello di evidenziare una dipendenza funzionale (anche perché tipicamente si misura una sola grandezza) ma di studiare come i valori misurati si distribuiscono e attorno a quali valori si addensano. Per analizzare graficamente queste caratteristiche si utilizza l'istogramma. Il primo passo nella costruzione di un istogramma è quello di dividere l'intervallo di variabilità della grandezza di cui si vuole l'istogramma in un certo numero di sotto-intervalli, i "*bin*", usualmente consecutivi e della stessa ampiezza[5]. Da un punto di vista più formale possiamo dire che l'istogramma è una rappresentazione grafica della distribuzione di dati numerici continui divisi in classi che vengono dette *bin*. L'istogramma appare come una successione di rettangoli contigui che hanno come base l'ampiezza del *bin* e un altezza proporzionale al numero di volte che il valore della grandezza è compreso tra i limiti del *bin*. Un criterio per la definizione dell'ampiezza dei *bin* è che ogni bin contenga un numero *congruo* di eventi. Come esempio supponiamo di avere ottenuto le seguenti $N = 20$ misure di una grandezza fisica (nelle opportune unità di misura): $2.36, 4.34, 2.03,$ $3.71, 4.28, 4.42,$ $3.75, 6.71, 6.53,$ $4.21, 6.98, 7.00, 0.87, 3.61, 5.31, 4.16, 7.31, 8.07, 5.78, 4.24$, e di volere analizzare graficamente come queste misure si distribuiscono. Come abbiamo già detto, iniziamo con il dividere l'intervallo di variabilità della grandezza in esame in *bin*. Una possibile scelta è quella di avere 9 *bin* tra i valori minimo e massimo delle misure, entrambi approssimati a valori ragionevolmente interizzati della grandezza. Il minimo è $x_{min} = 0.87$ approssimato a 0.0, il massimo è $x_{max} = 8.07$ approssimato a 9.0, ne segue che la larghezza del *bin* è $(9-0)/9 = 1$, misurato nelle stesse unità di x. Il passo successivo consiste nel contare il numero di volte che i valori sono contenuti in ognuno dei *bin* e riportare questo numero come altezza del corrispondente rettangolo come mostrato nella figura 4.8.

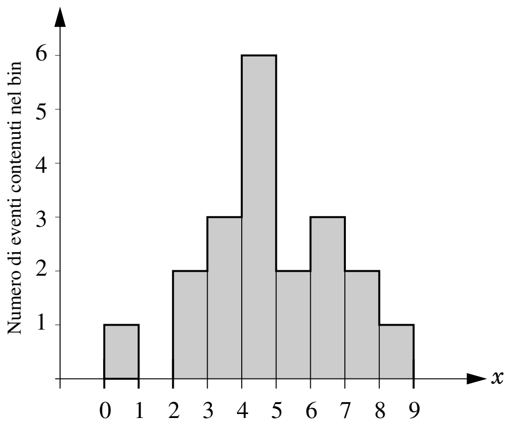


Figura 4.8: Esempio di istogramma dei 20 valori della variabile x indicati nel testo e raggruppati in 9 *bin*

Istogramma delle frequenze. Un altro modo di mostrare il contenuto dei *bin* è quello di dividerlo per il numero totale di eventi contenuti nell'istogramma. In questo caso si parla di istogramma delle frequenze. Se n_i è il contenuto del *bin* i-esimo, l'istogramma delle frequenze nell'i-esimo *bin* avrà il valore n_i/N.

L'espressione di istogramma è spesso riferita anche se in modo improprio a variabili casuali intere per le quali di dovrebbe parlare più correttamente di diagramma a barre. Come ad esempio consideriamo i valori usciti in un certo numero di lanci di un dado, come descritto nel seguente esempio. Lanciamo il dado $N = 200$ volte e annotiamo per ogni lancio il risultato. Contiamo quante volte è uscito 1, quante volte 2, ..., quante volte 6. Riportiamo questi conteggi ottenuti nell'asse delle ordinate disegnando una barra di altezza proporzionale al conteggio ottenuto, come mostrato nella figura 4.9. In questo modo abbiamo ottenuto l'istogramma o meglio il diagramma a barre degli N valori della variabile aleatoria "numero

[5]La uguale dimensione di *bin* non è obbligatoria anzi in certe circostanze, come vedremo nel seguito, è necessario utilizzare *bin* di dimensione differente

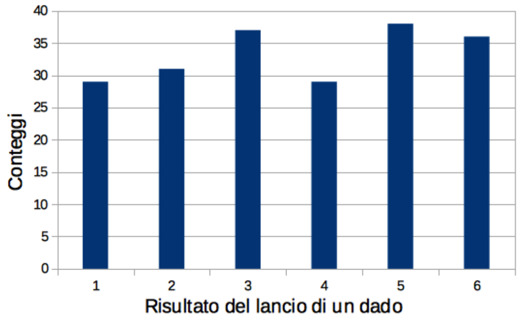


Figura 4.9: Istogramma dei numero usciti in 200 lanci di un dado. L'analisi qualitativa dell'istogramma mostra che il dado sembra non essere truccato.

uscito nel lancio del dado".

Sottolineiamo l'importanza dell'istogramma come strumento fondamentale per l'analisi dei dati di fisica; infatti anche attualmente importanti risultati di fisica sono annunciati mostrando degli istogrammi, come accadde nella scoperta del bosone di Higgs. Concludiamo questo capitolo mostrando nella figura 4.10 lo storico istogramma con il quale, nel 1974, è stata dimostrata l'esistenza del *charm* (il primo ad essere scoperto dei quark pesanti), creato nelle interazioni ad alta energia tra protoni.

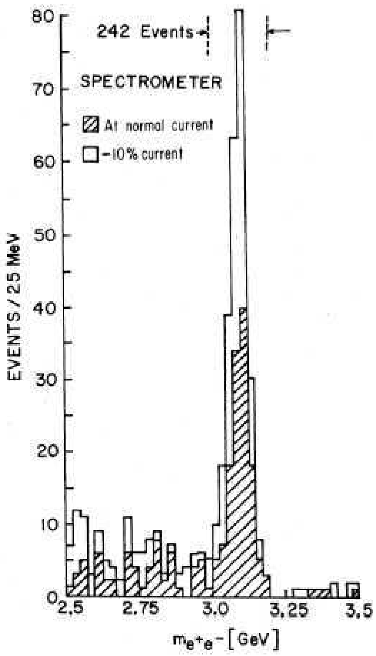


Figura 4.10: Istogramma originale che per la prima volta ha messo in evidenza l'esistenza del quark *charm* nelle interazioni $pp \rightarrow e^+e^-$. I due istogrammi sovrapposti servono a dimostrare che il picco non è dovuto ad un effetto strumentale