

Analisi Digitale dei Dati provenienti  
dal  
*Time Domain Sampler* (TDS)  
del satellite Wind.

Brando Bellazzini<sup>1</sup>  
Maggio 2002

---

<sup>1</sup>con la supervisione del Dott. L.Nocera dell'Istituto Processi Chimico Fisici-CNR.

# Indice

|  |    |
|--|----|
| Introduzione: Wind, Waves, e TDS       | 2  |
| 1 Smoothing e filtri digitali          | 5  |
| 2 Spettro: Data Windowing ed Overlap   | 9  |
| 3 Derivazione ed integrazione numerica | 14 |
| 4 Autocorrelazione                     | 19 |
| 5 Esponenti di Lyapunov e Caos         | 21 |
| 6 Conclusioni                          | 27 |
| Ringraziamenti                         | 30 |
| Bibliografia                           | 31 |
| Elenco delle figure                    | 32 |
| Elenco delle tabelle                   | 33 |

## Introduzione: Wind, Waves, e TDS

Wind Spacecraft, parte del programma ISTP (“International Solar Terrestrial-Physics”, di ESA, ISAS, NASA), è stato lanciato nel '94 per studiare il mezzo interplanetario e gli effetti delle perturbazioni, insite nel vento solare, sulla magnetosfera terrestre. Gli strumenti di Wind permettono di misurare campi magnetici ed elettrici, composizione e distribuzione energetica del plasma nel vento solare, raggi cosmici poco energetici, onde radio e onde plasma, e Gamma-Ray bursts. In particolare Waves Instrument, misura onde radio e onde plasma in un ampio range di frequenze (vedi fig. 4), la cui analisi, coordinata con le altre strumentazioni di Wind, vorrebbe fornire un'informazione completa dei fenomeni d'interazione del vento solare con la parte superiore della magnetosfera.

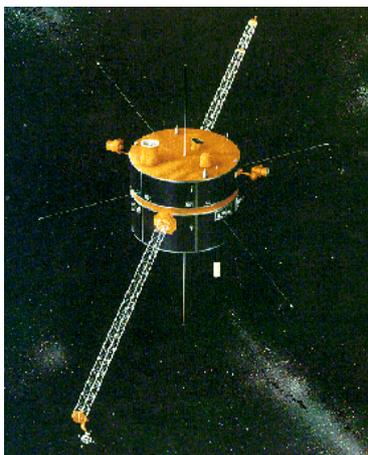


Figura 1: Il satellite Wind.

Time Domain Sampler (TDS), parte dell'esperimento Waves, campiona i campi elettromagnetici nel plasma con un tasso di campionamento che può raggiungere i 120000 punti al secondo (pts/s). Le forme d'onda elettrostatiche osservate con TDS sono essenzialmente di tre tipologie (fig. 3):

- pacchetti d'onda coerenti di Langmuir a frequenze  $f \simeq f_{pe}$  ( $f_{pe}$  frequenza di plasma elettronica)
- pacchetti d'onda coerenti con frequenze nel range  $f_{pi} \leq f \leq f_{pe}$  ( $f_{pi}$  frequenza di plasma ionica)
- picchi isolati non sinusoidali che durano meno di 1 ms: Isolated Electrostatic structures (IES)

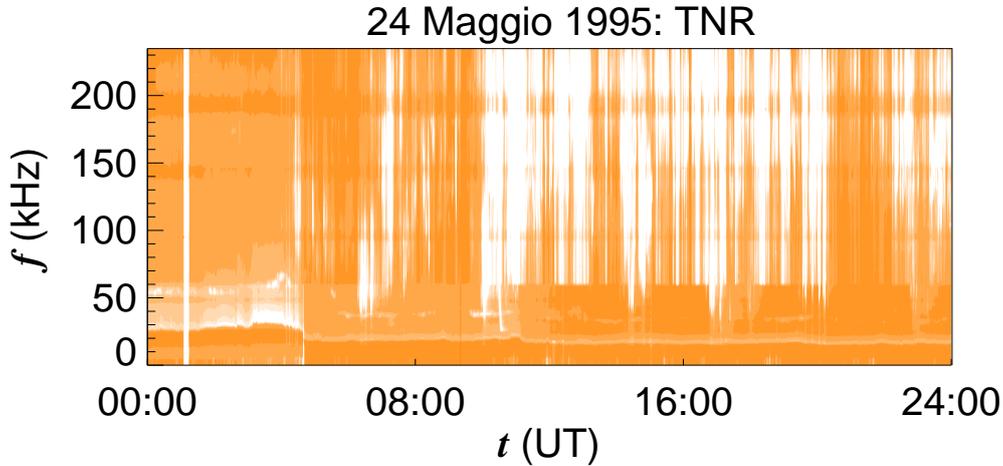


Figura 2: Esempio di spettro dinamico misurato da TNR, un altro strumento a bordo di Wind. L'intensità è misurata dalla tonalità del colore: a basse frequenze si nota la frequenza di plasma intorno a 20 KHz.

TDS effettua le misure delle forme d'onda grazie a due antenne ortogonali (per il campo elettrico) e ad un magnetometro a spire triassali (per il campo magnetico). Le due antenne sono un dipolo di  $2L_x = 100m$ , ed un dipolo molto più corto di  $2L_y = 15m$ . Per determinare il campo elettrico lungo  $x$  o  $y$  dividiamo la differenza di potenziale  $V$  per la corrispondente lunghezza efficace  $\bar{L}_x = 41.3m$  o  $\bar{L}_y = \frac{L_x}{10}$ , che differisce dalla lunghezza fisica dei dipoli a causa della presenza dello spacecraft. In generale il rapporto segnale-rumore è di un ordine di grandezza più piccolo per l'antenna più corta.

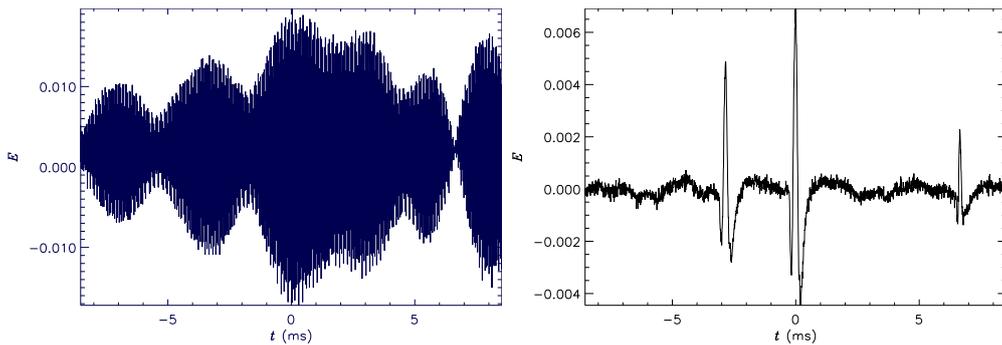


Figura 3: A sinistra pacchetto d'onda di Langmuir: la serie ha inizio alle 08:10:32.571(hh:mm:ss.ms) UT (tempo universale) del 96/02/02. A destra tre IES del 95/06/19 alle 09 : 17 : 39.761 .

Le forme d'onda elettriche che abbiamo considerato sono campionate a 120000 pts/s (quindi otteniamo informazioni spettrali per frequenze mino-

ri della frequenza di Nyquist  $f_c = 60\text{kHz}$ ), con una sensibilità di  $80\mu\text{V}$ , ed organizzate in serie temporali di 2048 punti per una durata complessiva di 17 ms.(vedi [1] oppure la pagina web <http://cdaweb.gsfc.nasa.gov> per maggiori dettagli).

I dati, forniti dall'osservatorio di Parigi–Meudon (relay center), sono stati acquisiti a tranches giornaliere (via FTP): la scrematura e la successiva selezione dei dati d'interesse è stata eseguita “a mano” con Interactive Data Language (IDL), pacchetto software grafico di largo impiego astrofisico.

Noi abbiamo eseguito un'analisi dei dati forniti da TDS, in particolare di IES, per mezzo di accurate tecniche di calcolo numerico di smoothing, spettro, integrazione, derivazione, autocorrelazione, esponenti di Lyapunov.

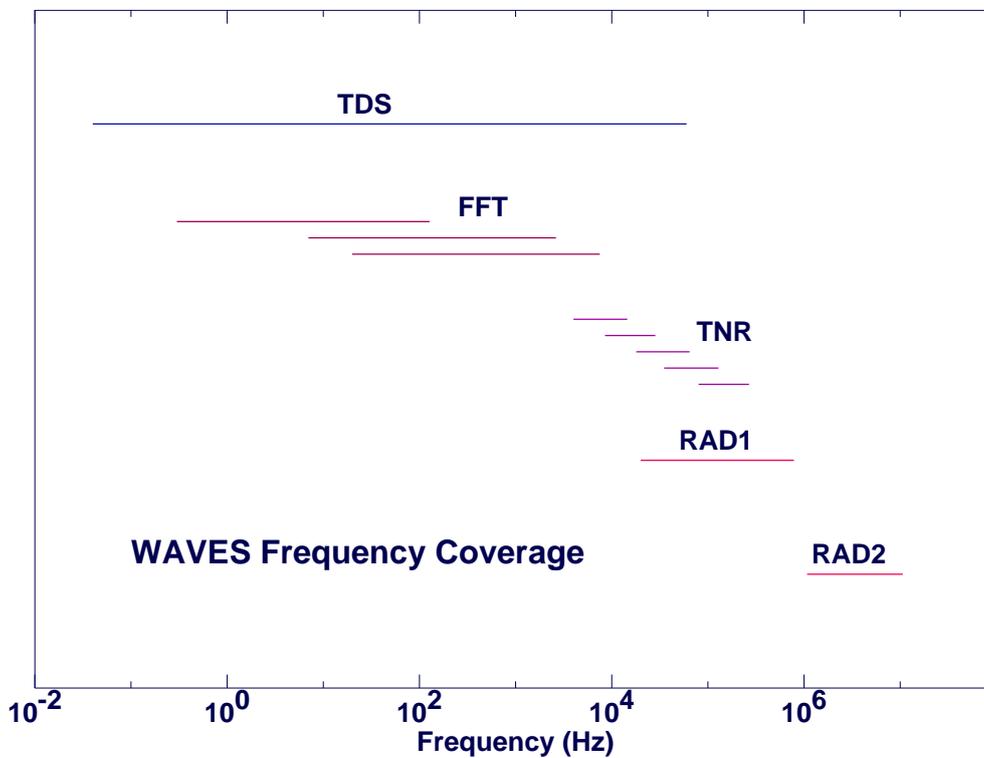


Figura 4: Diagramma sulla copertura spettrale degli strumenti di Wind.

# 1 Smoothing e filtri digitali

Un tipico filtro digitale di una serie temporale

$$f_i \equiv f(t_i) , t_i = t_o + i\Delta , \Delta = \text{passo del campionamento} \quad (1)$$

sostituisce ogni  $f_i$  con un'opportuna combinazione lineare di elementi della serie

$$g_i = \sum_{n=-N_l}^{N_r} c_n \cdot f_{i+n} , \quad (2)$$

dove  $N_l$  e  $N_r$  sono rispettivamente il numero di elementi della serie precedenti e successivi all' $i$ -esimo, per una finestra di valori  $[f_{i-N_l}, f_{i+N_r}]$ . Viene così prodotta una nuova serie temporale  $g_i$  che costituisce la serie filtrata.

La scelta più semplice dei coefficienti è prenderli uguali alla costante  $c_n = (N_l + N_r + 1)^{-1}$ , che significa porre  $g_i$  uguale alla media di  $f_i$  sulla finestra  $[f_{i-N_l}, f_{i+N_r}]$ . Il vantaggio di tale scelta, talvolta detta *Moving Window Averaging*, consiste nella proprietà di invarianza del momento zero (l'area sotto il grafico) e del momento primo (la media) se  $N_r = N_l$ , nella trasformazione  $f \rightarrow g$ . I limiti di questa scelta sono invece la modifica dei momenti di ordine superiore: ad esempio applicare un Moving Window Averaging nell'intorno di un massimo locale  $f_M$  comporta sicuramente la riduzione del massimo nella serie filtrata ovvero  $g_M \leq f_M$ , e intuitivamente l'appiattimento e l'allargamento della funzione  $g$  rispetto alla funzione  $f$  (vedi fig. 5 e 6).

Una scelta diversa dei coefficienti  $c_n$  che preservi i momenti di ordine superiore è l'idea base dei *Savitzky-Golay smoothing filters* [2] che abbiamo utilizzato per le forme d'onda che presentavano picchi isolati di notevole interesse (IES).

I Savitzky-Golay smoothing filters invece di utilizzare in ogni finestra un polinomio di grado zero (la costante) utilizzano polinomi di grado  $M$  (in generale  $M = 2$  o  $M = 4$ ): per ogni punto  $f_i$  si determina il polinomio  $p_i^M$  per mezzo del fit dei minimi quadrati sui valori della finestra  $[f_{i-N_l}, f_{i+N_r}]$ , e si pone

$$g_i = p_i^M(i) . \quad (3)$$

Il valore del polinomio viene calcolato solamente nel punto  $i$ -esimo; per il punto  $(i + 1)$ -esimo viene calcolato un nuovo polinomio  $p_{i+1}^M$ .

Anche se questo procedimento sembra molto dispendioso algebricamente, il fit dei minimi quadrati coinvolge solamente inversioni di matrici [2, cap15], e genera quindi i coefficienti polinomiali  $a_j$  come combinazione lineare delle  $f_k$  nella finestra  $[f_{i-N_l}, f_{i+N_r}]$ . Naturalmente questo assicura l'esistenza di un

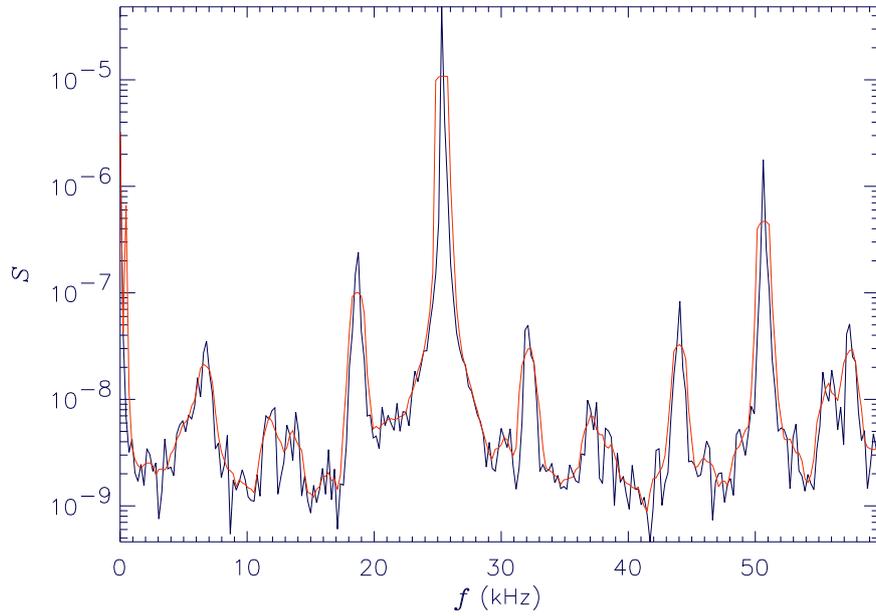


Figura 5: Moving Window Averaging applicato allo spettro del pacchetto d'onda di Langmuir di fig 3: è ben visibile l'allargamento e la riduzione dei picchi.

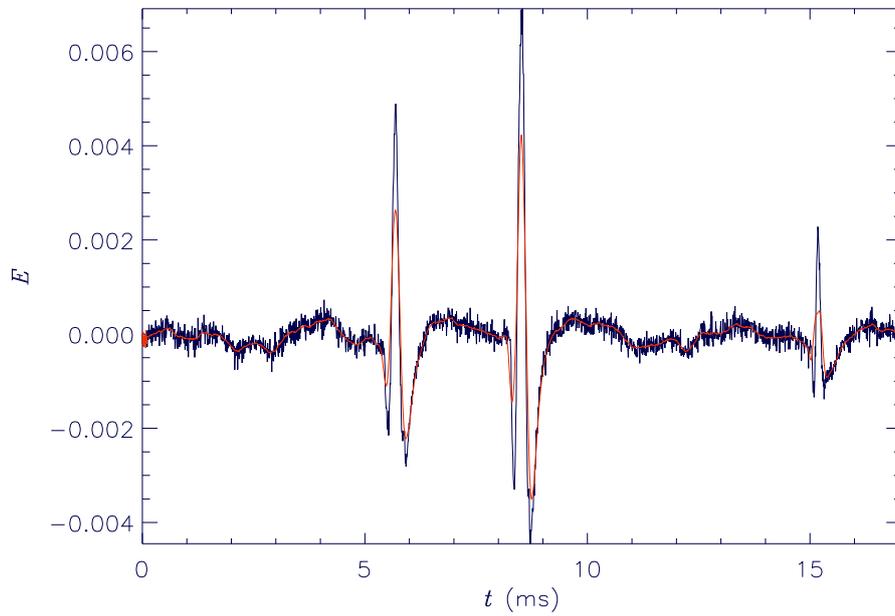


Figura 6: Moving Window Averaging alla serie di figura 3 con  $N_l = N_r = 25$  su tre IES: la riduzione dei picchi raggiunge il 50%.

insieme  $c_n$  per il quale la serie filtrata

$$g_i = p_i^M(i) = \sum_{k=0}^M c_k f_{n+i}, \quad (4)$$

viene formata mediante elementi adiacenti della serie originale (proprio come dichiarato nell'equazione 2), e “automaticamente” genera i coefficienti del polinomio soddisfacente il fit.

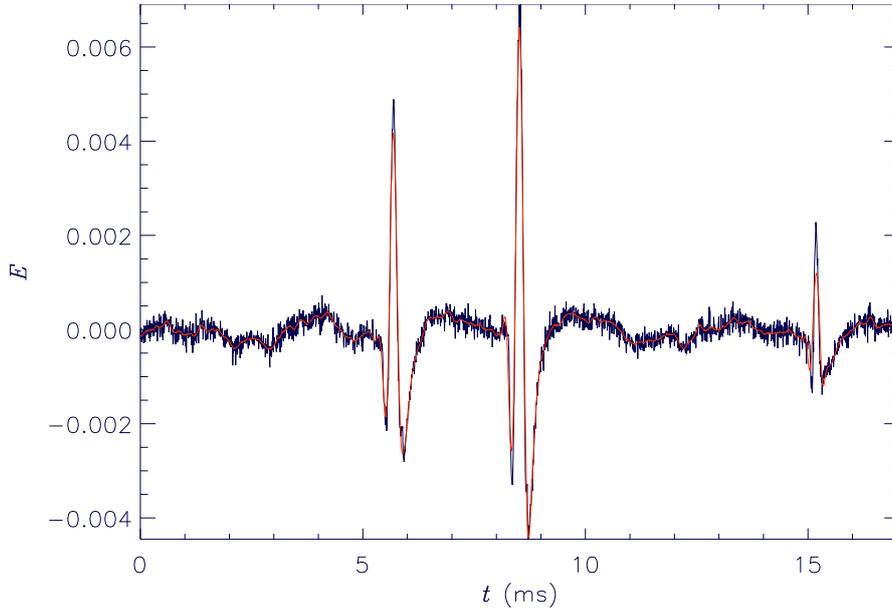


Figura 7: Savitzky-Golay smoothing filter eseguito con  $M = 4$  e  $N_l = N_r = 25$  sulla stessa serie della figura 6: la riduzione dei picchi è modesta.

Senza addentrarci nei calcoli del fit tra i punti  $f_{i-N_l}, f_{i-N_l+1}, \dots, f_i, \dots, f_{i+N_r}$  mediante il polinomio  $a_0 + a_1i + a_2i^2 + \dots + a_Mi^M$  (per dettagli vedi [2, cap14-15]), otteniamo le seguenti espressioni per i coefficienti  $c_n$  ed  $a_j$ :

$$c_n = \sum_{k=0}^M (\mathbf{A}^T \cdot \mathbf{A}^{-1})_{0k} \cdot n^k \quad \mathbf{a} = (\mathbf{A}^T \cdot \mathbf{A})^{-1} (\mathbf{A}^T \cdot \mathbf{f}), \quad (5)$$

dove  $\mathbf{A}_{jl} = j^l$ ;  $j = -N_l, \dots, N_r$ ;  $l = 0, \dots, M$ ;  $\mathbf{a}_l = a_l$ ;  $\mathbf{f}_k = f_k$ .

I risultati ottenuti con questo filtro digitale sono molto buoni come si vede nel confronto (fig. 6,7 e 8, 9) tra lo smoothing operato da IDL (che esegue un Moving Window Averaging) e dalla nostra routine (un Savitzky-Golay smoothing filter).

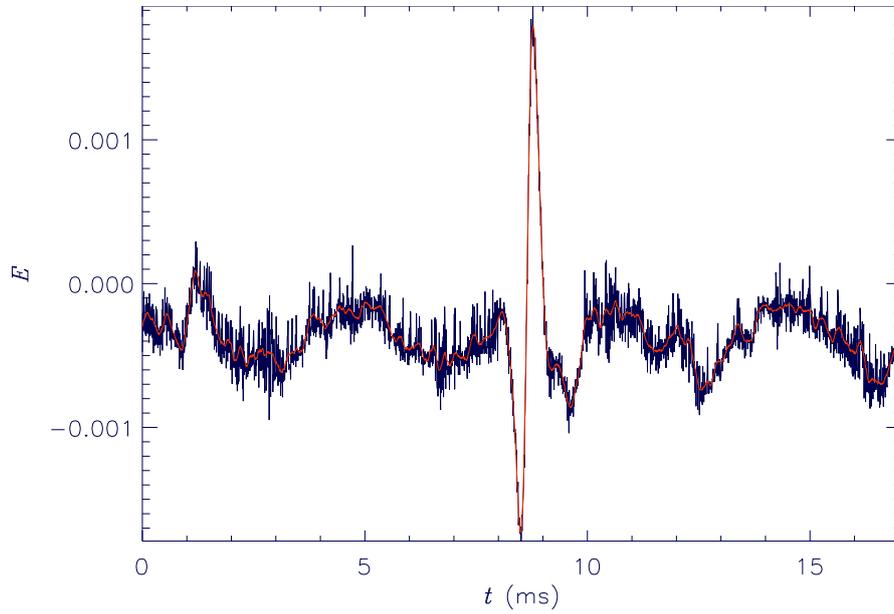


Figura 8: Savitzky-Golay smoothing filter eseguito con  $M = 4$  e  $N_l = N_r = 25$  su una IES del 96/02/02 delle 12:04:36.554 UT: si noti che il segnale è più “sporco” perchè misurato dall'antenna-Y di TDS.

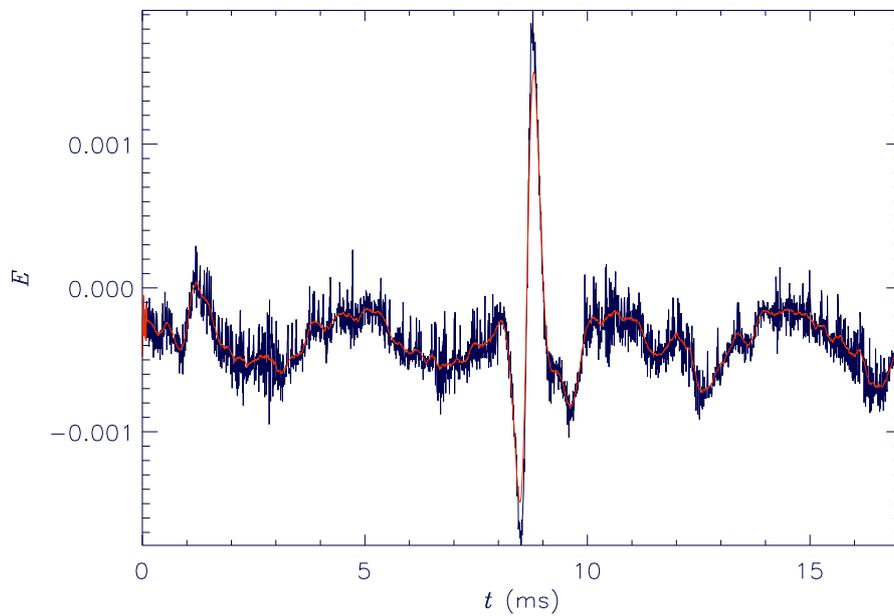


Figura 9: Moving Window Averaging con  $N_l = N_r = 25$  sulla stessa IES della figura 8.

## 2 Spettro: Data Windowing ed Overlap

Data una serie temporale

$$g_j = g(t_j), \quad t_j = t_0 + j\Delta, \quad j = 0, \dots, N-1 \quad (6)$$

possiamo calcolare la FFT (fast fourier transform)

$$G_k = \sum_{j=0}^{N-1} g_j e^{i2\pi \frac{jk}{N}}, \quad k = 0, \dots, N-1 \quad (7)$$

e stimare lo spettro di potenza  $P(f)$  di  $g$  grazie al periodogramma [2]

$$\begin{aligned} P_N(0) &= P(f_0) = \frac{1}{N^2} |G_0|^2 \\ P_N(f_k) &= \left( \frac{1}{N^2} |G_k|^2 + |G_{N-k}|^2 \right) \\ P_N(f_c) &= P(f_{N/2}) = \frac{1}{N^2} |C_{N/2}|^2, \end{aligned} \quad (8)$$

definito per le  $(n+1)$  frequenze

$$f_k = \frac{k}{N\Delta}, \quad k = 0, \dots, \frac{N}{2}. \quad (9)$$

In generale  $P_N(f_k) \neq P(f_k)$  ed anche  $P_\infty(f_k) \neq P(f_k)$ : infatti è possibile ricondurre  $P_N(f_k)$  alla media pesata di una certa funzione  $\bar{W}(f, f_k, N)$ , centrata al valore  $f_k$ , della funzione  $P(f)$  [2]

$$P_N(f_k) = \int_{-\infty}^{\infty} \bar{W}(f, f_k, N) P(f) df. \quad (10)$$

Per il periodogramma dato dalla eq. 8  $\bar{W}(f, f_k, N)$  è nota, e la sua espressione in termini dell' "offset"  $s = (f - f_k)N\Delta$  espresso in *bins* (cioè  $f = f_k \leftrightarrow s = 0$ ,  $f = f_{k+1} \leftrightarrow s = 1, \dots$ ) è la seguente

$$W(s, f_k, N) = \frac{1}{N^2} \left( \frac{\sin \pi s}{\sin \frac{\pi s}{N}} \right)^2 = \left| \sum_{k=0}^{N-1} e^{i2\pi \frac{sk}{N}} \right|^2. \quad (11)$$

$W(s, f_k, N)$  assume il valore massimo per  $s = 0$  (la frequenza  $f_k$  è la più "pesante") e decresce oscillando con un andamento  $\approx s^{-2}$ . Notiamo inoltre che, per  $s$  interi non nulli,  $W(s) = 0$  (le frequenze  $f_{j \neq k}$  hanno peso nullo).

Questo fenomeno di inquinamento da frequenze distanti alcuni bins prende il nome di *Leakage*. La tecnica, detta *Data Windowing*, che abbiamo

adottato per ridurre l'inquinamento da Leakage nasce dall'idea di modificare le equazioni 10 e 11 che specificano il legame e la dipendenza delle  $P_N(f_k)$  da  $P(f)$ . Tralasciando ancora una volta calcoli piuttosto tenici e laboriosi (per dettagli [2, cap13]), otteniamo, in luogo della 11, la seguente formula per  $W(s)$

$$W(s) = \frac{1}{W_{ss}} \left| \sum_{k=0}^{N-1} e^{i2\pi \frac{sk}{N}} w_k \right|^2 \approx \frac{1}{W_{ss}} \left| \int_{-N/2}^{N/2} \cos(2\pi \frac{sk}{N}) w(k - N/2) \right|^2 \quad (12)$$

$$W_{ss} = N \sum_{j=0}^{N-1} w_j^2,$$

dove  $w_j$  è un'opportuna *funzione finestra*. Le più comuni funzioni finestra, che conducono a risultati pressoché identici, sono le seguenti

$$w_j = 1 - \left| \frac{j - N/2}{N/2} \right| \quad \text{Barlett Window ,}$$

$$w_j = \frac{1}{2} \left( 1 - \cos \frac{2\pi j}{N} \right) \quad \text{Hann Window ,} \quad (13)$$

$$w_j = 1 - \frac{(j - N/2)^2}{N/2} \quad \text{Welch Window .}$$

L'effetto complessivo del Data Windowing è quindi modificare i “pesi” delle frequenze, diminuendo quello delle frequenze distanti ed aumentando quello delle frequenze nei bins adiacenti: questo porta ad un “ripulimento” dello spettro (fig. 10).

Un ulteriore aspetto che fin qui abbiamo tralasciato è la grande “indeterminazione” sulla stima della potenza spettrale a frequenza assegnata. Supponiamo ad esempio di campionare un segnale per un certo tempo  $T$  con un passo  $\Delta$ , di acquisire in tutto  $N$  punti e di calcolare il periodogramma per mezzo dell'equazione 8: come abbiamo già detto  $P_N(f_k) \neq P(f_k)$  ma intuitivamente potremmo attenderci un miglioramento della stima effettuata al crescere di  $N$ . In realtà la precisione non cresce affatto con il numero di punti campionati, anzi si può dimostrare che la varianza sul singolo valore del periodogramma alla frequenza  $f_k$  è uguale al quadrato del valore di aspettazione a quella stessa frequenza, ovvero che la deviazione standard è il 100% del valore.

L'informazione ulteriore che deriva da un campionamento più fitto (cioè da una diminuzione di  $\Delta$ ), o da un'acquisizione dati più lunga (cioè un aumento di  $N$  a  $\Delta$  costante) si traduce rispettivamente nell'innalzamento della frequenza di Nyquist  $f_c = \frac{1}{2\Delta}$  (l'informazione va dunque nelle nuove frequenze comprese tra la vecchia e la nuova  $f_c$ ), e nella risoluzione più fine

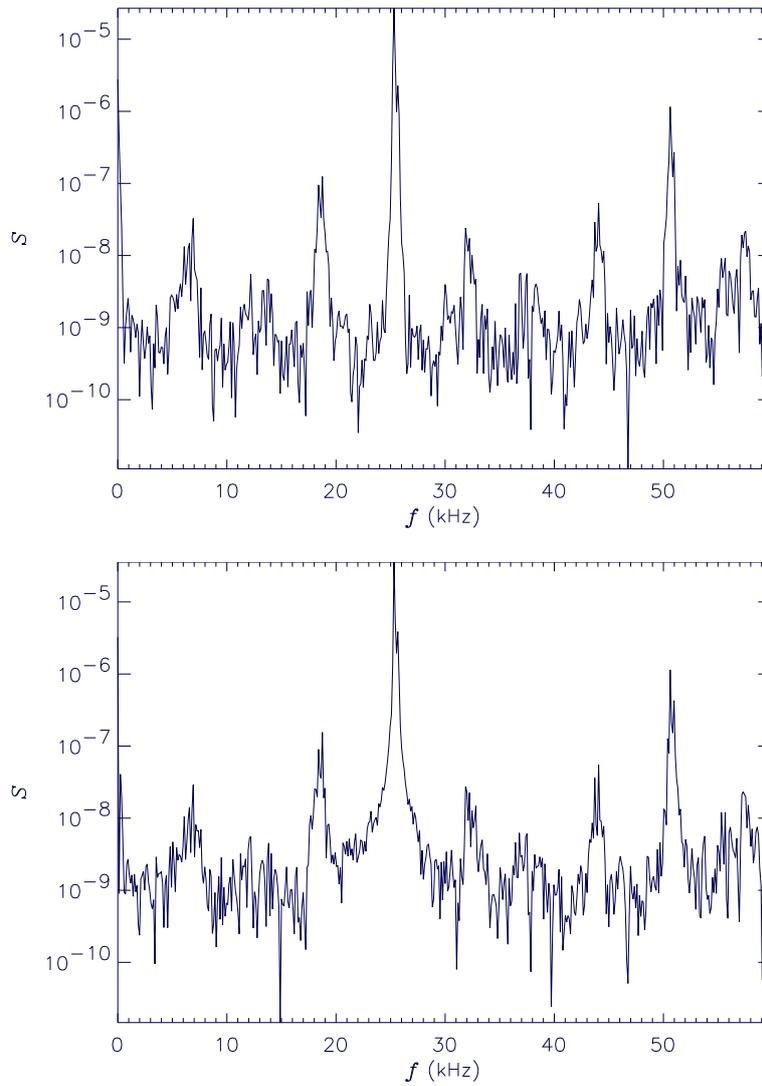


Figura 10: spettro del pacchetto d'onda di Langmuir di figura 3 eseguito con Welch Data Windowing (in alto) e senza Data Windowing (in basso).

delle frequenze discrete ( $f_{k+1} - f_k = \frac{\Delta}{N+M}$  contro il  $\frac{\Delta}{N}$  precedente) nel solito intervallo  $[0, f_c]$ .

La tecnica che abbiamo utilizzato per ridurre la varianza sulle stime fornite dal periodogramma, si basa sulla partizione dell'intervallo dati originale in  $k$  segmenti di  $2M$  punti ciascuno, e sul calcolo di  $k$  periodogrammi (tutti con le stesse  $M + 1$  frequenze discrete in  $[0, f_c]$ ): la media delle stime del potere spettrale ad ogni frequenza fornisce un nuovo periodogramma ed una nuova standard deviation ridotta, nell' ipotesi d' indipendenza statistica tra i  $k$  segmenti, di un fattore  $\sqrt{k}$  (ogni segmento per quanto già detto presenta varianza costante indipendentemente dal numero di punti).

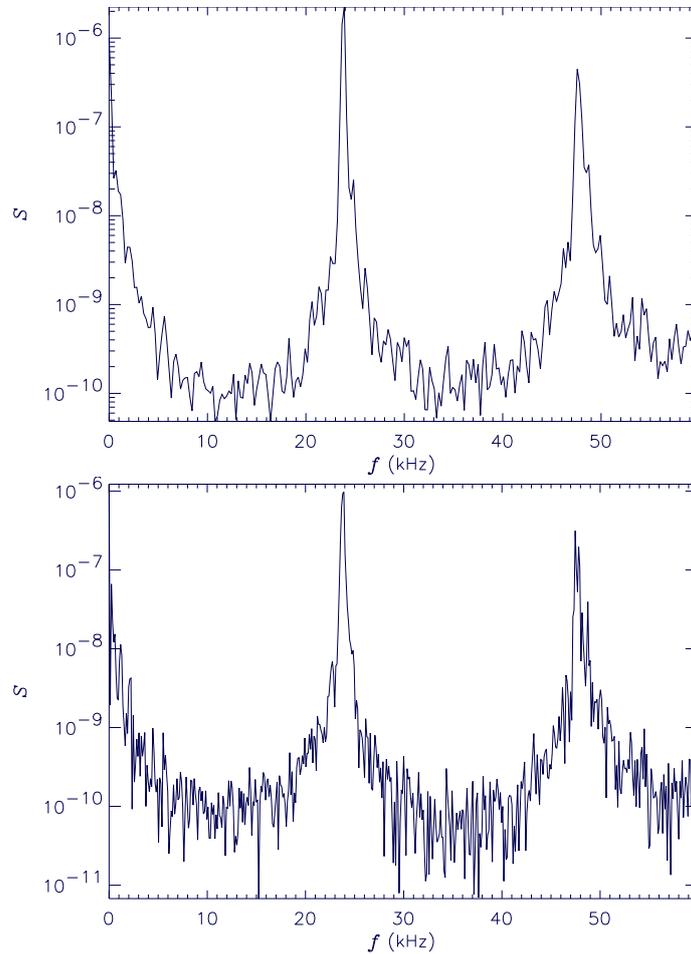


Figura 11: spettro del pacchetto d'onda di Langmuir di figura 16 eseguito con overlap (in alto) e senza overlap (in basso).

L'ipotesi d'indipendenza statistica si rileva tuttavia troppo ristretta per i segnali che vogliamo studiare, i quali presentano un'autocorrelazione non

trascurabile. Per ovviare a questo problema effettuiamo il cosiddetto *Overlap*, cioè la divisione dell'intervallo originale in  $k$  segmenti non disgiunti che condividono parte dei punti degli intervalli adiacenti, in maniera da lasciare in ciascuno di essi "memoria" di quello precedente. Il guadagno in deviazione standard risulta ottimale, o quasi ottimale, con la condivisione di mezzo intervallo, che permette la riduzione della varianza a  $\frac{11}{9k}$  del valore di aspettazione. Anche la tecnica della segmentazione e dell'overlap portano ad un ripulimento dello spettro come si vede in fig. 11.

Pur sorvolando sull'importanza di un'accurata analisi spettrale dei segnali, sottolineiamo la differenza tra gli spettri di IES e pacchetti d'onda coerenti di Langmuir: i primi vanno a zero piuttosto rapidamente senza presentare picchi notevoli o eccessivamente marcati, i secondi invece mostrano una frequenza dominante ed i suoi multipli interi, dovuti alle interazioni non lineari tra onde.

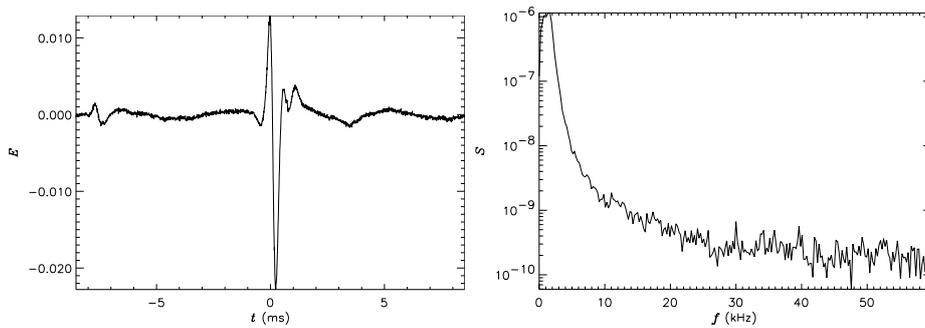


Figura 12: IES del 96/02/02 alle 12:04:36.554 UT (a sinistra), e spettro associato (a destra).

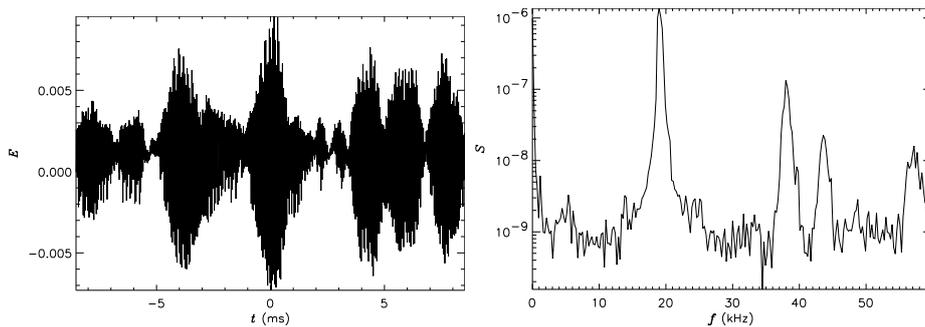


Figura 13: pacchetto d'onda di Langmuir del 95/05/24 alle 17:53:57.630 UT (a sinistra), e spettro associato (a destra).

### 3 Derivazione ed integrazione numerica

Il problema che vogliamo affrontare in questo paragrafo riguarda le approssimazioni numeriche che si effettuano per calcolare derivata ed integrale di una funzione, nota la serie prodotta dal suo campionamento. Nel nostro caso la serie

$$f_k = f(t_k) , t_k = t_0 + k\Delta , k = 1, \dots, N \quad (14)$$

è prodotta dal campionamento delle differenze di potenziale tra l'estremo dell'antenna  $x$  e lo spacecraft.

Facciamo presente inoltre che per le onde analizzate in questa relazione il campo elettrico è allineato alla direzione di flusso del vento solare, e che in un sistema di riferimento solidale al plasma si osservano solamente campi elettrostatici unidirezionali [1]: questo comporta che il campo elettrico si sposti rigidamente nello spazio, trasportato dal vento (con velocità  $v$  compresa tra 300 e 600km/s), e che l'integrazione o la derivazione temporale sia essenzialmente equivalente a quelle spaziale<sup>2</sup>:

$$\frac{d}{dx} = \frac{1}{v} \frac{d}{dt} . \quad (15)$$

La serie misurata da TDS (delle differenze di potenziale) costituisce anche una misura diretta del campo elettrico che, poichè la lunghezza d'onda è molto grande rispetto alla lunghezza dell'antenna (per dettagli [1]), è dato semplicemente dalla seguente equazione

$$E_{x_k} = \frac{f_k}{\bar{L}_x \cdot \cos \theta} . \quad (16)$$

L'interesse fisico per la serie derivata è giustificato dalla prima equazione di Maxwell, infatti per quanto detto sopra

$$\frac{\partial E_y}{\partial y} = \frac{\partial E_z}{\partial z} = 0 , \quad \frac{\partial E_x}{\partial x} = 4\pi\rho = -4\pi e(n_e - n_i) , \quad (17)$$

con  $n_e$  e  $n_i$  densità elettronica e ionica. La serie integrale del campo elettrico è interessante invece perchè fornisce, a meno del segno, il potenziale  $\phi$ :

$$-\frac{\partial \phi}{\partial x} = E_x , \quad -\frac{\partial \phi}{\partial y} = 0 , \quad -\frac{\partial \phi}{\partial z} = 0 . \quad (18)$$

---

<sup>2</sup>C'è da dire però che l'antenna  $x$  del satellite non è in generale allineata con la direzione del campo elettrico, e quindi è sempre presente un fattore numerico compreso tra -1 ed 1 a moltiplicare, come effetto della proiezione del campo sull'antenna stessa. Per conoscere l'angolo  $\theta$  formato tra antenna e campo elettrico all'istante di interesse, bisogna interpolare misurazioni di *spin* dello spacecraft che sono effettuate ogni 32 secondi. Sottolineiamo inoltre che per la durata della misura (17ms) l'angolo può essere considerato costante, essendo la frequenza di spin di Wind  $2.0820 \pm 5 \cdot 10^{-5}$ rad/s.

Come punto di partenza per il calcolo delle derivate mostriamo un classico esempio di approssimazione numerica: direttamente dalla definizione di derivata possiamo scrivere

$$\dot{f}(t_0) = \frac{f(t_0 + \Delta) - f(t_0)}{\Delta} + \sigma(\Delta), \quad (19)$$

dove  $\sigma(\Delta)$ , l' "errore", è un infinitesimo in  $\Delta$ ; oppure, equivalentemente,

$$\dot{f}(t_0) = \frac{f(t_0 + \Delta) - f(t_0 - \Delta)}{2\Delta} + \bar{\sigma}(\Delta). \quad (20)$$

Ciò che non è equivalente nelle due espressioni è l'ordine di infinitesimo dell'errore:  $\bar{\sigma}$  è del 1° ordine in  $\Delta$ , infatti sviluppando  $f(t_0 \pm \Delta)$  in serie di Taylor (supporremo verificate tutte le ipotesi di regolarità della  $f(t)$  che saranno necessarie) si ottiene

$$f(t_0 \pm \Delta) = f(t_0) \pm \Delta \dot{f}(t_0) + \Delta^2 \ddot{f}(t_0) + o(\Delta^2) \quad (21)$$

che implica  $\bar{\sigma} = o(\Delta)$  a differenza di  $\sigma(\Delta)$  che in generale è solamente un infinitesimo in  $\Delta$ . Naturalmente questo ci convince a impiegare la 20 come prima approssimazione di  $\dot{f}(t_0)$  visto che l'errore tende a zero più rapidamente.

In effetti, più in generale, data la serie temporale dell'eq. 14 è possibile esprimere la derivata di  $f$  in  $t_k$  come combinazione lineare di  $n$  elementi vicini della serie

$$\begin{aligned} \dot{f}_k &= \sum_{j=-N_l}^{N_r} a_j f_{k+j} + E_{N_r, N_l, k}(\Delta) \quad \text{se } (N_l + 1) \leq k \leq (N - N_r), \\ \dot{f}_k &= \sum_{j=0}^n a_{k,j} f_{k+j} + E_{N_r, N_l, k}(\Delta) \quad \text{se } k \leq (N_l + 1), \\ \dot{f}_k &= \sum_{j=0}^n a_{k,j} f_{k-j} + E_{N_r, N_l, k}(\Delta) \quad \text{se } k \geq (N_r), \end{aligned} \quad (22)$$

dove  $E_{N_r, N_l, k}(\Delta)$  è infinitesimo di ordine  $n$ . Una forma più espressiva dell'equazione 22 si ottiene utilizzando la notazione matriciale

$$\dot{\mathbf{f}} = \mathbf{A} \cdot \mathbf{f} + \mathbf{E}(\Delta), \quad (23)$$

dove  $\mathbf{A}$  è una matrice non singolare con zeri quasi ovunque, eccetto che in una banda centrata sulla diagonale (vedi tab. 1).

La scelta dei coefficienti dipende inoltre dalla natura del fenomeno in esame: per i problemi di propagazione ondosa risultano molto efficaci scelte "sbilanciate in avanti" (cioè con  $N_r \geq N_l$ ), ed in particolare la scelta a nove

$$\mathbf{A} = \begin{pmatrix} b_{1,1} & b_{1,2} & b_{1,3} & 0 & \cdots & & & \cdots & & 0 \\ b_1 & b_2 & b_3 & 0 & \cdots & & & & & \vdots \\ 0 & b_1 & b_2 & b_3 & 0 & \cdots & & & & \\ 0 & 0 & b_1 & b_2 & b_3 & 0 & \cdots & & & \\ \vdots & \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & & & \\ & & & \ddots & \ddots & \ddots & \ddots & \ddots & & \\ & & & & \cdots & 0 & b_1 & b_2 & b_3 & 0 & 0 \\ & & & & & \cdots & 0 & b_1 & b_2 & b_3 & 0 \\ \vdots & & & & & & \cdots & 0 & b_1 & b_2 & b_3 \\ 0 & \cdots & & & & & \cdots & 0 & b_{N,N-2} & b_{N,N-1} & b_{N,N} \end{pmatrix} \quad (24)$$

Tabella 1: Esempio di matrice  $\mathbf{A}$  con  $N_r = N_l = 1$

| k                | $a_{-3}$ | $a_{-2}$ | $a_{-1}$ | $a_0$   | $a_1$   | $a_2$   | $a_3$   | $a_4$   | $a_5$  |
|------------------|----------|----------|----------|---------|---------|---------|---------|---------|--------|
| 1                | -109584  | 322560   | -564480  | 752640  | -705600 | 451584  | -188160 | 46080   | -5040  |
| 2                | -5040    | -64224   | 141120   | -141120 | 117600  | -70560  | 28224   | -6720   | 720    |
| 3                | 720      | -11520   | -38304   | 80640   | -50400  | 26880   | -10080  | 2304    | -240   |
| $4 \div (N - 5)$ | -240     | 2880     | -20160   | -18144  | 50400   | -20160  | 6720    | -1440   | 144    |
| N-4              | 144      | -1536    | 8064     | -32256  | 0       | 32256   | -8064   | 1536    | -144   |
| N-3              | -144     | 1440     | -6720    | 20160   | -50400  | 18144   | 20160   | -2880   | 240    |
| N-2              | 240      | -2304    | 10080    | -26880  | 50400   | -80640  | 38304   | 11520   | -720   |
| N-1              | -720     | 6720     | -28224   | 70560   | -117600 | 141120  | -141120 | 64224   | 5040   |
| N                | 5040     | -46080   | 188160   | -451584 | 705600  | -752640 | 564480  | -322560 | 109584 |

Tabella 2: tabella dei coefficienti che abbiamo utilizzato:  $N_r = 5$ ,  $N_l = 3$ ,  $n = 9$ .

punti con  $N_r = 5$  e  $N_l = 3$  (vedi tab. 2). Questo stesso schema di approssimazione numerica della derivazione per mezzo della matrice  $\mathbf{A}$  “a nove punti sbilanciata in avanti” fornisce buoni risultati per la risoluzione dell’equazione integro–differenziale di Vlasov nei plasmi [4].

Per ottenere i coefficienti dell’equazione 22, o equivalentemente gli elementi della matrice  $\mathbf{A}$ , interpoliamo gli  $n = N_r + N_l + 1$  punti con polinomi di Lagrange di grado  $n - 1$

$$f(t) = \sum_{k=1}^n f_k P_k(t) \quad \text{con} \quad P_k(t) = \prod_{m \neq k} \frac{t - t_m}{t_k - t_m} = \sum_{j=1}^n B_{kj} t^{j-1} \quad (25)$$

ed approssimiamo  $\dot{f}(t_k)$  con il valore della derivata del polinomio  $\sum f_k P_k(t)$  calcolato in  $t_k$ : otteniamo così una combinazione lineare delle  $f_k$  corretta entro termini di ordine  $o(\Delta^n)$ .

I coefficienti  $B_{kj}$  si ottengono facilmente considerando che

$$P_k(t_i) = \delta_{ki} = \sum_{j=1}^n B_{kj} t_i^{j-1} \quad (26)$$

cioè che  $B_{kj}$  è la matrice inversa della matrice  $C_{ij} = t_i^{j-1}$  (matrice di Vandermonde).

Esempi di derivazione numerica di serie temporali fornite da TDS sono mostrate nelle fig. 14 e 15.

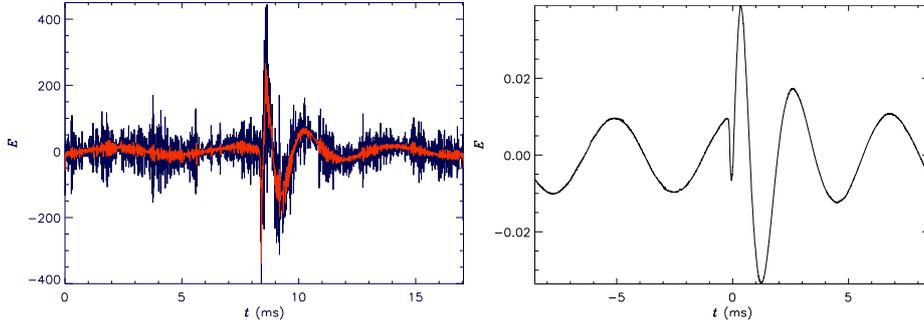


Figura 14: derivata (a sinistra) di una IES del 96/02/02 alle 13:13:43.546 UT (a destra): in rosso eseguita da IDL, in blu da noi con lo schema a nove punti della tab. 2.

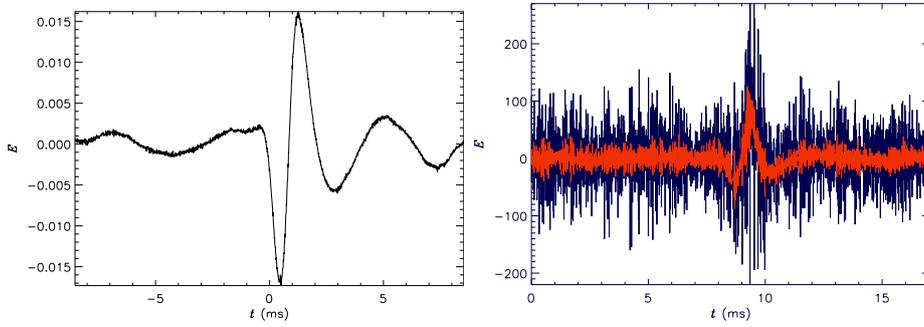


Figura 15: derivata (a sinistra) di una IES del 97/09/16 alle 11:23:35.996 UT (a destra): in rosso eseguita da IDL, in blu da noi con lo schema a nove punti della tab. 2.

Il problema inverso di risalire dalla  $f_k$  alla serie integrale  $\varphi_k$ , e conseguentemente dalla serie del campo elettrico  $E_{x_k}$  alla serie del potenziale  $\phi_k$ , si risolve discretizzando l'equazione differenziale 18 per mezzo della matrice  $\mathbf{A}$  di derivazione già ricavata, ottenendo così un sistema lineare di  $N$  equazioni in  $N$  incognite

$$\sum_{k=1}^N \mathbf{A}_{jk} \varphi_k = f_j \quad (27)$$

$$\sum_{k=1}^N \mathbf{A}_{jk} \phi_k = -E_{x_j} \quad (28)$$

Nel nostro caso le incognite sono  $N=2048$  e la matrice  $\mathbf{A}$  ha dimensione  $2048 \cdot 2048$ . Per aumentare l'efficienza nella risoluzione del sistema abbiamo usato la *decomposizione LU* [2] della matrice  $\mathbf{A}$ .

Esempi di integrazione numerica delle serie temporali fornite da TDS sono riportati nelle fig. 16 e 17. Riportiamo per inciso il fatto che ogni serie numerica di cui abbiamo eseguito l'integrazione presentava una serie integrale dall'andamento analogo a quello mostrato dalle figure; questo fatto è probabilmente riconducibile alla presenza di un campo elettrico medio costante che conferisce una pendenza media costante ai grafici del potenziale.

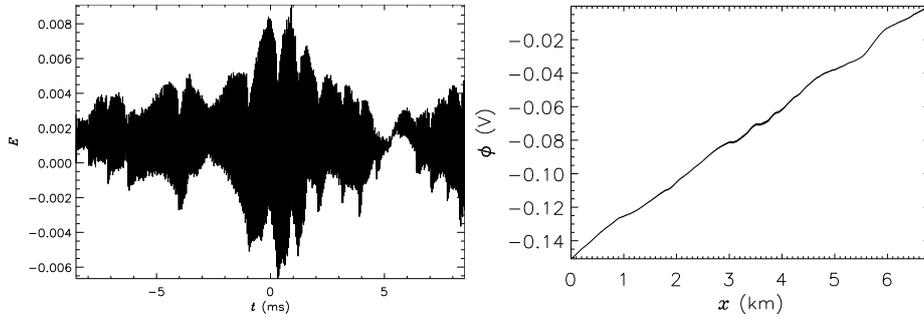


Figura 16: Pacchetto di Langmuir del 96/02/02 alle 15:50:13.660 UT (a sinistra), e suo integrale (a destra): è ben visibile una pendenza media costante nella serie integrale.

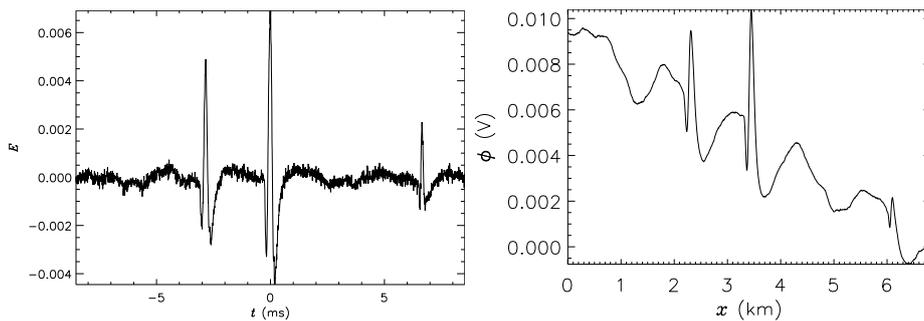


Figura 17: Come per la serie di fig. 16, ma per l'IES di fig. 3.

## 4 Autocorrelazione

La correlazione tra due serie temporali ad  $N$  punti

$$g_k = g(t_k), f_k = f(t_k), t_k = t_0 + k \cdot \Delta, k = 0, \dots, N - 1 \quad (29)$$

costituisce anch'essa una serie temporale definita dall'espressione

$$[Corr(g, f)]_j = \sum_{k=0}^{N-1} g_{j+k} f_k, j = 0, \dots, N - 1. \quad (30)$$

L' autocorrelazione non è altro che la correlazione di una serie con se stessa.

I punti che devono essere calcolati sono  $N$ , e per ognuno di essi si devono eseguire  $N$  somme, per un totale di  $N^2$  operazioni. Per cercare di ridurre il carico computazionale a grandi  $N$  sfruttiamo il *Teorema di correlazione*, il quale afferma che

$$F(Corr(g, f)) = F(g)F^*(f), \quad (31)$$

dove  $F$  indica la trasformata discreta di Fourier e l'asterisco la coniugazione complessa. Per eseguire i calcoli delle  $F$  e di  $F^{-1}$  per risalire alla correlazione abbiamo utilizzato due FFT (fast fourier transform) e una  $FFT^{-1}$ , ciascuna con un dispendio di  $N \ln_2 N$  operazioni. Il un carico complessivo computazionale risulta dunque anch'esso dell'ordine di  $N \ln_2 N$ , un valore estremamente ridotto rispetto a  $N^2$  e quindi molto più vantaggioso (per maggiori dettagli [2, cap12-13]).

Esempi di autocorrelazione dei segnali forniti da TDS sono mostrati nelle fig. 18-21.

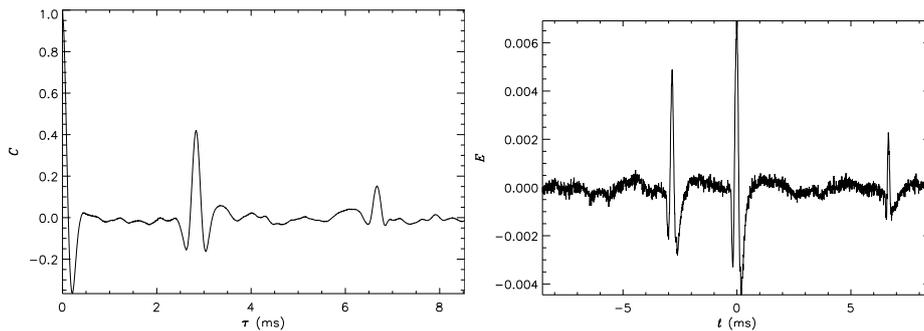


Figura 18: autocorrelazione (a sinistra) dell'IES della fig. 3 (a destra).

E' ben visibile dalle figure il contrasto tra l'autocorrelazione di un IES e di un pacchetto d'onda coerente di Langmuir.

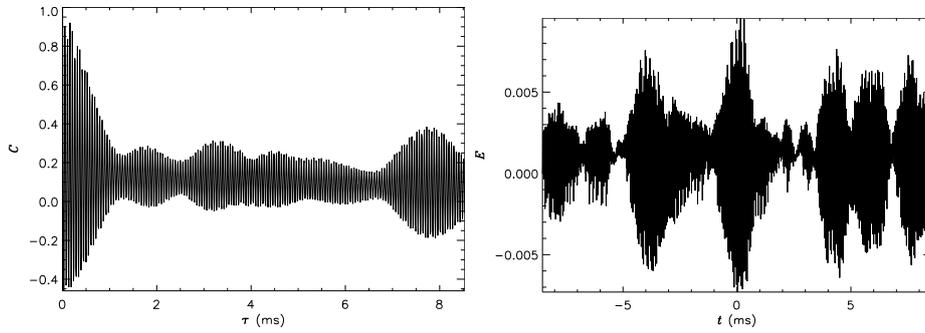


Figura 19: autocorrelazione (a sinistra) del pacchetto di Langmuir della fig. 13 (a destra).

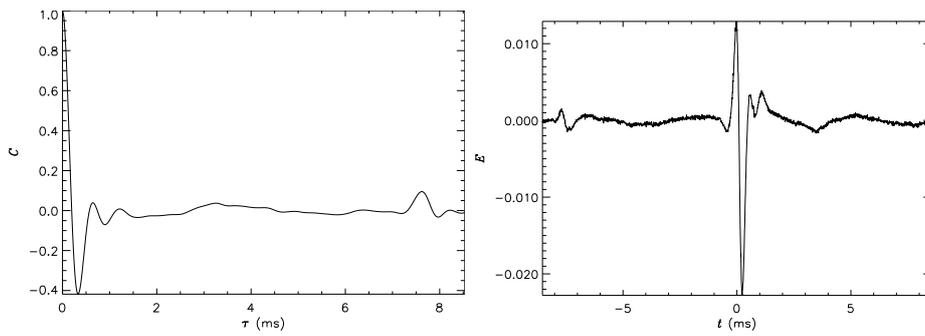


Figura 20: autocorrelazione (a sinistra) dell'IES della fig. 12 (a destra).

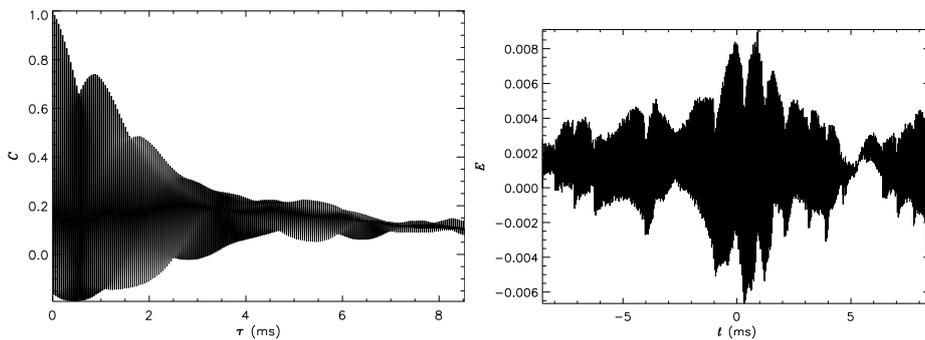


Figura 21: autocorrelazione (a sinistra) del pacchetto di Langmuir della fig. 16 (a destra).

## 5 Esponenti di Lyapunov e Caos

Gli esponenti di Lyapunov sono le principali quantità usate in dinamica non lineare per distinguere i comportamenti caotici dai non caotici, per misurare cioè la divergenza o la convergenza di stati inizialmente molto vicini, per stimare il periodo di tempo oltre il quale le predizioni diventano impossibili. Per definizione un sistema è caotico se presenta almeno un'esponente positivo.

Il calcolo numerico degli esponenti di Lyapunov di un sistema dinamico

$$\dot{\mathbf{x}}(t) = \mathbf{f}(\mathbf{x}, t), \quad \mathbf{x}(t) = (x^1(t), x^2(t), \dots, x^d(t)) \in \mathbf{R}^d \quad (32)$$

si basa su algoritmi di algebra lineare [8] che si applicano quando siano note le  $d$  serie temporali

$$x^1(t_i), x^2(t_i), \dots, x^d(t_i) \quad t_i = (i - 1)\Delta \quad (33)$$

oppure, equivalentemente, le  $(d - 1)$  derivate rispetto al tempo

$$x^1(t_i), x^{1'}(t_i), x^{1''}(t_i), \dots, x^{1^{(d)}}(t_i) \quad (34)$$

di una componente (della prima ad esempio) del vettore  $\mathbf{x}$ . Nel caso sperimentale non possediamo  $d$  serie, ma una solamente  $s(t_i) = s^i$ , frutto delle misurazioni di una qualche grandezza fisica, e ci chiediamo se è possibile ricostruire un opportuno sistema dinamico, cioè se possiamo ricavare le  $d - 1$  serie derivate

$$s^i, s^{i'}, s^{i''}, \dots, s^{i^{(d)}} \quad (35)$$

dalle quali poi eseguire il calcolo degli esponenti. Naturalmente non sappiamo neanche qual è il  $d$  buono; procederemo quindi empiricamente “vedendo ciò che accade” al variare di  $d$ . Per determinare le  $s^{i^{(j)}}$ , calcoliamo le derivate numeriche mediante differenze finite con punti distanti  $n_d$  da  $s^i$  (che chiameremo punti  $n_d$ -vicini).

$$\begin{aligned} s^{i'} &\simeq \frac{s^{i+n_d} - s^i}{n_d \Delta}, \\ s^{i''} &\simeq \frac{s^{i+2n_d} - 2s^{i+n_d} + s^i}{(n_d \Delta)^2}, \\ &\vdots \end{aligned} \quad (36)$$

Intuitivamente il miglior valore per  $n_d$  sembrerebbe  $n_d = 1$ , cioè il più piccolo possibile, tuttavia vi sono indicazioni che, per tenere conto della omogeneità statistica della serie  $s^i$ , è opportuno scegliere  $n_d$  maggiore della distanza a

cui cade il primo minimo della funzione di auto-correlazione di  $s^i$  [7]. Anche in questo caso adotteremo il metodo empirico “vedendo ciò che accade” per  $n_d$  diversi. Tecnicamente  $n_d$  prende il nome di *stride* o *tempo di ritardo*.

Una volta determinate tutte le derivate fino all'ordine  $d$ , possiamo rappresentarle in uno spazio  $d$ -dimensionale che chiameremo *spazio delle fasi ricostruito* della serie  $s^i$ . Operiamo ora la seguente semplificazione: le derivate ottenute dall'equazione 36 sono combinazioni lineari dei punti  $n_d$ -vicini  $s^i, s^{i+n_d}, \dots, s^{i+(d-1)n_d}$ , e quindi si può passare dallo spazio delle fasi allo spazio dei punti  $n_d$ -vicini mediante una trasformazione lineare. Da un punto di vista computazionale è effettivamente più comodo lavorare in questo nuovo spazio nel quale la serie  $s^i$  è detta *immersa*. Lo spazio dei punti  $n_d$ -vicini prende perciò il nome di *spazio d'immersione*  $S^d$  della serie  $s^i$ . Analogamente  $d$  prende il nome di *dimensione d'immersione*.

Il passo successivo alla ricostruzione dei punti  $\mathbf{x}^i = (s^i, s^{i+n_d}, \dots, s^{i+(d-1)n_d})$  nello spazio d'immersione, è l'approssimazione del flusso incognito

$$\Phi^{t_e} : \mathbf{x}^i \rightarrow \mathbf{x}^{i+n_e} \quad t_e = n_e \Delta : \text{tempo di evoluzione} \quad (37)$$

per mezzo di un'opportuna funzione interpolante  $\phi : S^d \rightarrow S^d$  le cui componenti  $l$ -esime sono date dalla sovrapposizione di  $m$  funzioni di base  $X_k(\mathbf{x})$

$$\mathbf{y}_l = [\phi(\mathbf{x})]_l = \sum_{k=1}^m \phi_{kl} X_k(\mathbf{x}). \quad (38)$$

Per determinare i coefficienti  $\phi_{kl}$ , assegnate le  $X_k(\mathbf{x})$ , è possibile prendere due strade: imporre che il flusso approssimato passi per le immagini  $\mathbf{y}^k = \Phi^{t_e}(\mathbf{z}^k)$  di punti assegnati  $\mathbf{z}^k \in S^d$  con  $k = 1, \dots, m$ , oppure utilizzare il metodo dei minimi quadrati su questi stessi punti. La seconda scelta riduce l'effetto del rumore insito nei dati osservativi [5]. I punti  $\mathbf{z}^k$  che entrano in gioco nel calcolo dei coefficienti  $\phi_{kl}$ , sono alcuni degli evoluti vicini a  $\mathbf{x}^i$ , verificanti delle condizioni speciali affinché la soluzione rispetto ai coefficienti dell'equazione 38 sia effettivamente possibile. Ad esempio se scegliessimo di imporre la prima delle due condizioni, i coefficienti della componente  $l$ -esima sarebbero dati dal sistema

$$\mathbf{y}_l = \begin{pmatrix} y_l^1 \\ \vdots \\ y_l^m \end{pmatrix} = \begin{pmatrix} \phi_l(\mathbf{z}^1) \\ \vdots \\ \phi_l(\mathbf{z}^m) \end{pmatrix} = \begin{pmatrix} \sum_{k=1}^m \phi_{kl} X_k(\mathbf{z}^1) \\ \vdots \\ \sum_{k=1}^m \phi_{kl} X_k(\mathbf{z}^m) \end{pmatrix}, \quad (39)$$

ovvero dal seguente sistema di equazioni lineari

$$\mathbf{Y} = \begin{pmatrix} y_1^1 & \cdots & y_d^1 \\ \vdots & \ddots & \vdots \\ y_1^m & \cdots & y_d^m \end{pmatrix} = \begin{pmatrix} X_1(\mathbf{z}^1) & \cdots & X_m(\mathbf{z}^1) \\ \vdots & \ddots & \vdots \\ X_1(\mathbf{z}^m) & \cdots & X_m(\mathbf{z}^m) \end{pmatrix} \cdot \begin{pmatrix} \phi_{11} & \cdots & \phi_{1d} \\ \vdots & \ddots & \vdots \\ \phi_{m1} & \cdots & \phi_{md} \end{pmatrix} \quad (40)$$

il quale è risolvibile se la matrice

$$\mathbf{A} = \begin{pmatrix} X_1(\mathbf{z}^1) & \cdots & X_m(\mathbf{z}^1) \\ \vdots & \ddots & \vdots \\ X_1(\mathbf{z}^m) & \cdots & X_m(\mathbf{z}^m) \end{pmatrix} \quad (41)$$

è non singolare. In pratica ai fini del calcolo effettivo dei  $\phi_{kl}$  e per il buon funzionamento degli algoritmi di calcolo degli esponenti di Lyapunov, è importante che la matrice  $\mathbf{A}$  sia sufficientemente “lontana” dalla condizione di singolarità.

Se i punti  $\mathbf{z}^k$  sono distinti, la scelta di *funzioni radiali*

$$X_k(\mathbf{x}) = \sqrt{r^2 + \|\mathbf{x} - \mathbf{z}^k\|^2} \quad (42)$$

come funzioni di base garantisce l’invertibilità della matrice  $\mathbf{A}$ . Una buona scelta per il *parametro di stiffness*  $r$  è renderlo dipendente dal punto  $\mathbf{x}^i$  con la sostituzione  $r_i = r \cdot \max_k \|\mathbf{x}^i - \mathbf{z}^k\|_\infty$  ( $\|\cdot\|_\infty$  è la norma del max).

In questo modo, alla fine, ci siamo ricostruiti il flusso  $\phi$ , la cui matrice Jacobiana trasposta al punto  $\mathbf{x}^i$  è dato da

$$D\phi^{tr}(\mathbf{x}^i) = \begin{pmatrix} \frac{\partial X_1(\mathbf{x}^i)}{\partial x_1} & \cdots & \frac{\partial X_m(\mathbf{x}^i)}{\partial x_1} \\ \vdots & \ddots & \vdots \\ \frac{\partial X_1(\mathbf{x}^i)}{\partial x_d} & \cdots & \frac{\partial X_m(\mathbf{x}^i)}{\partial x_d} \end{pmatrix} \cdot \begin{pmatrix} \phi_{11} & \cdots & \phi_{1d} \\ \vdots & \ddots & \vdots \\ \phi_{m1} & \cdots & \phi_{md} \end{pmatrix} = \mathbf{D} \cdot \mathbf{C} . \quad (43)$$

Adesso siamo in grado di calcolarci gli esponenti di Lyapunov  $\lambda_j$  (dalla definizione stessa) grazie alla conoscenza della successione di jacobiani del flusso approssimato: il metodo più efficiente è fattorizzare la matrice

$$D\phi^p(\mathbf{x}^1) \equiv \prod_{i=1}^p D\phi(\mathbf{x}^i) \quad (44)$$

come il prodotto di una matrice ortogonale  $Q^p$  e  $p$  matrici triangolari superiori  $R^l$  ad elementi positivi sulla diagonale, da cui si ottiene

$$D\phi^p(\mathbf{x}^1) = Q^p \cdot R^p \cdot \dots \cdot R^1 : \quad (45)$$

gli esponenti di Lyapunov sono forniti infine dal limite [8]

$$\lambda_j = \lim_{p \rightarrow +\infty} \lambda_j^p = \lim_{p \rightarrow +\infty} \frac{1}{pt_e} \sum_{i=1}^p \ln_2 R_{jj}^i \quad j = 1, \dots, d . \quad (46)$$

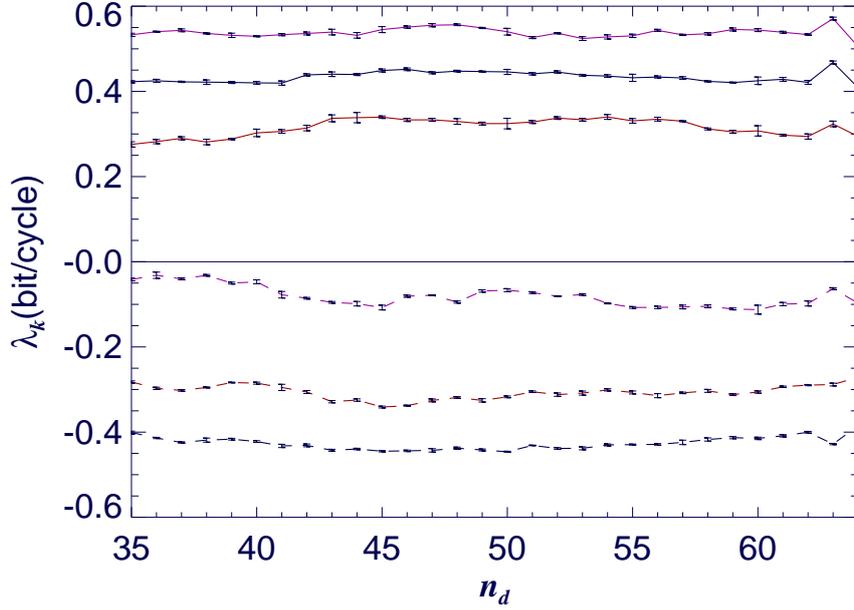


Figura 22: esponenti di Lyapunov (con errori) del pacchetto di Langmuir di fig 16. Con linea tratteggiata sono riportati anche gli esponenti calcolati per la serie temporale invertita: come si vede un esponente è spurio, e gli altri sono positivi, quindi la serie è caotica.

In pratica per valutare questo limite si effettua la media, cui resta associata anche una varianza, di tutte le osservazioni di  $\lambda_j^p$  per  $n - n_0 \leq p \leq n$  con  $n = N/n_e$  e  $n_0 = n/10$ :

$$\bar{\lambda}_j = \sum_{p=n-n_0}^n \lambda_j^p / n_0, \quad (47)$$

$$\sigma_j^2 = \sum_{p=n-n_0}^n (\lambda_j^p - \bar{\lambda}_j)^2 / n_0. \quad (48)$$

Il procedimento che abbiamo eseguito per il calcolo degli esponenti di Lyapunov ha il notevole inconveniente di generarne di spuri, non legati cioè alla dinamica del sistema che genera il sistema, ma determinati dalle scelte arbitrarie di cui l'algoritmo necessita, in particolare della dimensione di immersione  $d$ .

Per ovviare a questo problema abbiamo utilizzato il metodo di Parlitz [5], che permette di discriminare esponenti spuri e non, e di stabilire se la dimensione d'immersione scelta è troppo elevata: consideriamo la serie temporale  $h^i$  ricavata dalla serie  $s^i$  per inversione temporale,

$$h^i = s^{N-i+1} \quad i = 1, \dots, N \quad (49)$$

e calcoliamo nuovamente gli esponenti di Lyapunov della serie  $h^i$ . Gli esponenti “veri” della serie  $s^i$  cambieranno segno (intuitivamente se due punti divergevano sotto l’azione del flusso, nell’inversione temporale convergono), al contrario di quelli fittizi che, essendo legati all’algoritmo e non alle proprietà dinamiche del sistema, cambieranno in maniera differente. Queste considerazioni ci permettono effettivamente di capire se  $d$  deve essere diminuito, e quali sono gli esponenti di Lyapunov genuini.

Un ulteriore criterio di bontà delle scelte arbitrarie effettuate nell’esecuzione dell’algoritmo ( $n_d, n_e, m, \dots$ ), e quindi degli esponenti di Lyapunov prodotti, è fornito dal parametro di controllo *errore di predizione*: per  $n = N/n_e$  volte operiamo con il flusso  $\phi$  sulla nuvola degli  $\mathbf{z}^k$ , e calcoliamo quindi  $e_{pred}$  secondo la seguente formula

$$e_{pred}^2 = \sum_{i=1}^n \sum_{k=1}^d [\phi_k(\mathbf{x}^i) - x_k^{i+n_e}]^2 / (nd). \quad (50)$$

I nostri risultati saranno considerati validi e consistenti per piccoli valori (rispetto agli esponenti) dell’errore di predizione.

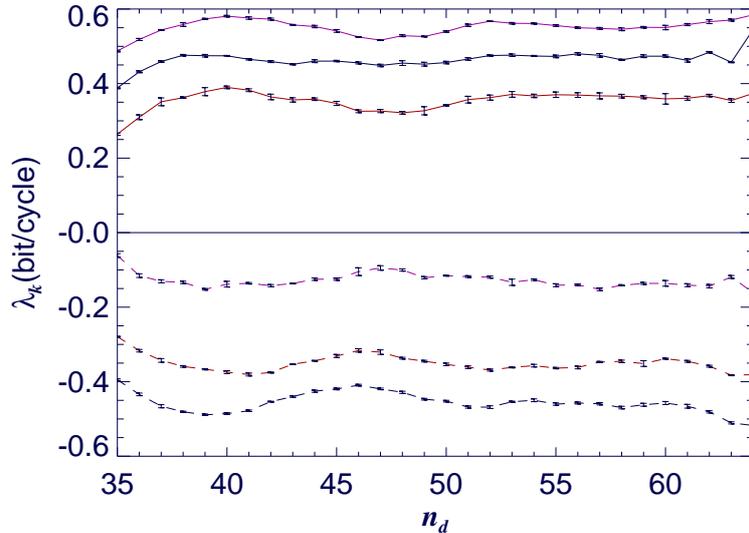


Figura 23: esponenti di Lyapunov (ed errori) di un altro pacchetto di Langmuir del 96/02/02 alle 22:37:41.881 UT. Con linea tratteggiata sono riportati anche gli esponenti calcolati per la serie temporale invertita: anche questa ne ha uno spurio, e due positivi, quindi è caotica.

L’unità di misura degli esponenti di Lyapunov è il bit/(unità di tempo). Per capire l’origine di questa notazione bisogna pensare che un esponente di Lyapunov positivo comporta una crescita esponenziale della distanza tra due punti dello spazio delle fasi inizialmente vicini. Questo fatto implica che un

cubetto di tali punti accrescerà esponenzialmente il proprio volume, e quindi anche il numero di stati possibili del sistema dinamico caotico.

Secondo la definizione di Shannon [10] un sistema che può occupare un numero maggiore di stati può fornire più informazione, e pertanto un sistema dinamico caotico *crea informazione* al passare del tempo. Sempre secondo Shannon [10] il logaritmo in base 2 del numero di stati di un sistema definisce la misura dell'informazione che si dirà espressa in *bits* (binary digits).

Applicando questi concetti al nostro caso troviamo che il volume occupabile dal nostro sistema al tempo  $t$  è  $\mathcal{V} \sim e^{\lambda t}$  ( $\lambda$  esponente di Lyapunov), e che pertanto l'informazione da esso esprimibile secondo Shannon è

$$\mathcal{J} = \log_2 \mathcal{V} \text{ bits} = \lambda t \text{ bits} . \quad (51)$$

È evidente che questo volume diviene sempre più grande al variare di  $t$ , e risulta quindi più significativo il tasso di creazione di informazione per unità di tempo:

$$\frac{d\mathcal{J}}{dt} = \lambda \text{ bits}/(\text{unità di tempo}) . \quad (52)$$

Se eseguiamo il calcolo degli esponenti che abbiamo illustrato per le serie temporali fornite da TDS, otteniamo effettivamente degli esponenti positivi (vedi fig. 22, 23 e 24) che evidenziano la natura caotica delle serie temporali. Tanto pacchetti d'onda di Langmuir quanto IES presentano (solamente) esponenti positivi, anche se probabilmente quelli riscontrati per gli IES, sono generati più dal rumore, o comunque dalla piccola componente alla frequenza di plasma, che non dalla portante stessa (in questo caso il picco).

Nelle figure qui riportate, l'unità di misura degli esponenti, poichè sono stati normalizzati alle rispettive frequenze di plasma, è bits/ciclo.

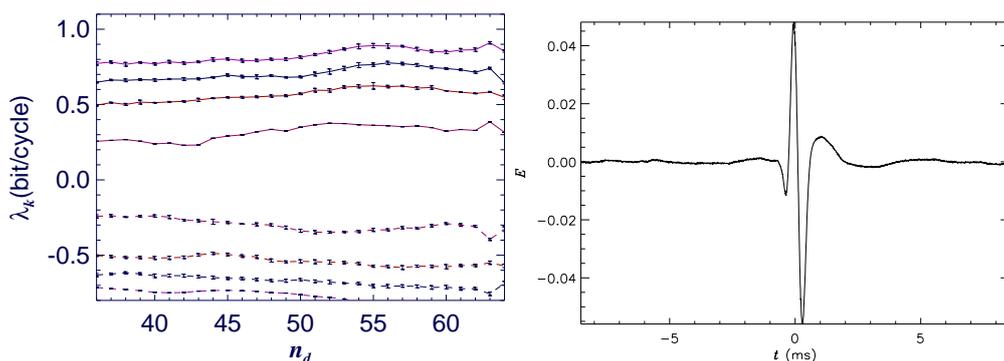


Figura 24: esponenti di Lyapunov (in alto) di una IES del 96/02/02 alle 11:38:30.714 UT (in basso). Con linea tratteggiata sono riportati anche gli esponenti calcolati per la serie temporale invertita. Come si osserva gli esponenti sono tutti puri e positivi, quindi la serie è caotica.

## 6 Conclusioni

In questa relazione sono state studiate le serie temporali di campi elettrici campionate dal *Time Domain Sampler* a bordo del satellite *Wind* nel *Vento Solare* vicino. Lo scopo della relazione era duplice:

1. mostrare le tecniche di elaborazione numerica e digitale che sono state utilizzate per l'analisi delle serie temporali fornite da misurazioni dei campi elettrici nel vento solare
2. riportare i risultati più importanti ottenuti con questa metodologia d'indagine.

### Tecniche di analisi dati

I principali strumenti di calcolo numerico e di analisi dati che abbiamo utilizzato sono: *Savitzky-Golay smoothing filters*, *fast fourier transform*, *data windowing*, *overlap*, derivazione e integrazione per mezzo del *metodo delle differenze finite spostate in avanti*, calcolo dell'autocorrelazione, *calcolo degli esponenti di Lyapunov*.

Molte di queste tecniche fanno parte di pacchetti software di largo impiego scientifico, che puntano ad una vasta utenza e quindi alla generalità d'impiego, che tuttavia si "paga" in termini di efficienza di calcolo ed accuratezza. Pertanto tutte le nostre analisi sono invece basate su specifiche routine scritte in Fortran.

Come primo esempio dei benefici ottenuti con l'impiego di routine specifiche consideriamo il caso dello smoothing. L'utilizzo del filtro operato dal pacchetto software IDL, il quale impiega un *Moving Window Averaging* che notoriamente ha il limite di ridurre ed allargare i picchi, produce un generale appiattimento (vedi ad esempio fig. 5 e 6). Nella nostra analisi abbiamo preferito utilizzare *Savitzky-Golay smoothing filters* [2] i cui benefici sono illustrati nelle fig. 7 ed 8.

Un secondo tipo di analisi coinvolge gli spettri di potenza calcolati tramite FFT. L'applicazione della sola FFT comporta i noti inconvenienti di *data leakage* tra bins di frequenza diversi, e di una grande varianza [2] sulla determinazione del singolo valore dello spettro. Abbiamo affrontato questi due problemi rispettivamente con le tecniche di *Data Windowing* ed *Overlap*, i cui effetti sono illustrati dalle fig. 10 e 11.

Un'ulteriore analisi è stata necessaria per il calcolo della distribuzione spaziale della densità di carica elettronica a partire dai dati primari forniti da *Wind*. Tale quantità è stata ricavata per derivazione numerica, come

prescritto dalla prima equazione di Maxwell (vedi eq.17). La tecnica di derivazione che abbiamo usato è quella *delle differenze finite spostate in avanti* [3] che riduce il problema della ricerca della serie derivata al calcolo degli elementi di un'opportuna matrice  $\mathbf{A}$  che operi sulle serie di partenza commettendo un errore di ordine assegnato. Nel nostro caso abbiamo usato uno schema di ordine nove il cui errore è  $o(\Delta^9)$ , con  $\Delta$  passo del campionamento temporale. Gli elementi di queste matrici sono noti e tabulati [3].

Strettamente correlato al calcolo della derivata, è la determinazione del potenziale elettrico che si ottiene per integrazione del campo elettrico secondo l'eq. 18. La tecnica impiegata sostanzialmente effettua l'inversione della matrice  $\mathbf{A}$  di derivazione, per mezzo di speciali *decomposizioni LU* che rendono più efficiente maneggiare matrici di  $2048 \cdot 2048$  elementi. La bontà di queste integrazioni è stata inoltre verificata dalla derivazione effettuata da IDL sulle serie integrali trovate, ottenendo un ottimo accordo con le serie temporali di partenza.

La forte irregolarità delle serie osservate ha suggerito la necessità di eseguire il calcolo delle funzioni di autocorrelazione (fig. 18, 19, 20, 21). Tuttavia questa quantità fornisce solamente una valutazione qualitativa dell'incoerenza dei segnali sperimentali. Negli ultimi vent'anni l'Analisi Matematica ha sviluppato algoritmi che consentono di estrarre informazioni quantitative dalle serie temporali. Uno di questi algoritmi più moderni è certamente il calcolo degli esponenti di Lyapunov [8].

Il calcolo di questi esponenti per le serie provenienti da TDS, costituisce l'aspetto più originale della relazione, perchè si qualifica come la parte non standard dell'analisi effettuata (ed effettivamente si è trattato di un'opera quasi "artigianale" caratterizzata da aspetti euristici ed intuitivi, anche se rigorosamente verificati, tipici di un vero e proprio esperimento). A causa della relativa brevità delle serie temporali (2048 pts) si è reso necessario l'utilizzo di un algoritmo molto sofisticato [5] che permette di aggirare i noti vincoli sulla *dimensione dello spazio d'immersione* [9]. I risultati di un tale esperimento sono riassunti dalle fig. 22, 23 e 24.

## Risultati principali

Benché la presente relazione si indirizzi principalmente allo studio di tecniche di analisi dati avanzate, non vanno trascurati alcuni risultati originali.

Un risultato notevole ottenuto mediante l'integrazione numerica è il particolare andamento dei potenziali (fig.16 e 17) il quale fa pensare alla presenza di un campo elettrico medio costante nel vento solare. Questo è in linea con risultati analoghi ricavati presso l'osservatorio di Meudon (A. Mangeney, comunicazione privata).

L'altro risultato importante è la caratterizzazione delle serie temporali fornite da Wind come serie caotiche, cioè provenienti da processi dinamici con almeno un esponente di Lyapunov positivo (vedi fig. 22, 23 e 24).

L'aspetto teorico–modellistico, che dovrebbe associare le strutture osservate dal satellite a soluzioni dell'equazioni costitutive del plasma nel vento solare, non è stato qui affrontato per ragioni di tempo e spazio, ma costituisce un aspetto complementare all'analisi svolta.

## Ringraziamenti

- Istituto per i Processi Chimico–Fisici – CNR, Pisa: L. Nocera, routines IDL, routines per gli esponenti di Lyapunov
- Observatoire de Paris–Meudon, France: A. Mangeney WIND co-investigator, routines di acquisizione dati WIND-TDS
- Space Science Laboratory, Univ. California Berkeley, USA: C. Salem, telemetria WIND
- Univ. Minnesota at Minneapolis: P. Kellogg, WIND-TDS co-investigator: autorizzazione acquisizione dati TDS
- Goddard Space Flight Center, MD, USA: M. Kaiser, WIND co-principal investigator, dat su CDAWeb
- Personale CDAWeb: US mirror: dati sui parametri chiave di WIND

## Riferimenti bibliografici

- [1] A. Mangeney, C. Salem, J. L. Bougeret, C. Perche, R. Manning, P. J. Kellogg, K. Goetz, S. J. Mounson, J. -M. Bosqued *Wind observation of coherent electrostatic waves in the solar wind*. *Annales Geophysicae* **17**, 307-320, 1999..
- [2] William H. Press, William T. Vetterling, Saul A. Teukolsky, Brian P. Flannery *Numerical Recipes in Fortran: The Art of scientific computing*. Cambridge University Press, 1992.
- [3] W. G. Bickley *Formulae for numerical differentiation*. The Mathematical Gazette, 1939.
- [4] Z. Sedláček, L. Nocera *Linear Vlasov Plasma Oscillation in the Fourier Transformed Velocity Space*. *Journal of Physics A*, 2002, in press.
- [5] U. Parlitz *Identification of true and spurious Lyapunov exponents from time series*. *International Journal of Bifurcation and Chaos*, Vol. **2**, 155-165, 1992.
- [6] L. Nocera, G. Bertin, S. V. Bulanov, F. Pegoraro *Chaotic dynamics of magnetic X-points*. ISPP-17 “Piero Caldirola”, *Theory of fusion Plasmas*, 411-416, 1996.
- [7] H. Isliker, A. O. Benz *Non-Linear properties of the dynamics of bursts and flares in the solar and stellar coronae*. *Astronomy and Astrophysics* **285**, 663-674, 1994.
- [8] G. Benettin, L. Galgani, A. Giorgilli, J.-M. Strelcyn *Lyapunov characteristic exponents for smooth dynamical systems and for hamiltonian systems; a method for computing all of them. Part 1: Theory, Part 2: Numerical Application*. *Meccanica* **15**, 9-30, 1980.
- [9] J.-P. Heckmann, D. Ruelle *Fundamental limitations for estimating dimension and Lyapunov exponents in dynamical system*. *Physica D*, **56**, 185-187, 1992.
- [10] C. E. Shannon *A Mathematical Theory of Communication*. The Bell System Technical Journal, Vol. **27**, 379-423, 623-656, 1948.

## Elenco delle figure

|    |   |    |
|----|---|----|
| 1  | Il satellite Wind. . . . .  | 2  |
| 2  | Esempio di spettro dinamico misurato da TNR, un altro strumento a bordo di Wind.<br>L'intensità è misurata dalla tonalità del colore: a basse frequenze si nota la frequenza<br>di plasma intorno a 20 KHz. . . . .   | 3  |
| 3  | A sinistra pacchetto d'onda di Langmuir: la serie ha inizio alle 08:10:32.571(hh:mm:ss.ms) UT<br>(tempo universale) del 96/02/02. A destra tre IES del 95/06/19 alle 09 : 17 : 39.761 . . . . .   | 3  |
| 4  | Diagramma sulla copertura spettrale degli strumenti di Wind. . . . .  | 4  |
| 5  | Moving Window Averaging applicato allo spettro del pacchetto d'onda di Langmuir<br>di fig 3: è ben visibile l'allargamento e la riduzione dei picchi. . . . .   | 6  |
| 6  | Moving Window Averaging alla serie di figura 3 con $N_l = N_r = 25$ su tre IES: la<br>riduzione dei picchi raggiunge il 50%. . . . .  | 6  |
| 7  | Savitzky-Golay smoothing filter eseguito con $M = 4$ e $N_l = N_r = 25$ sulla stessa serie<br>della figura 6: la riduzione dei picchi è modesta. . . . .  | 7  |
| 8  | Savitzky-Golay smoothing filter eseguito con $M = 4$ e $N_l = N_r = 25$ su una IES del<br>96/02/02 delle 12:04:36.554 UT: si noti che il segnale è più "sporco" perchè misurato<br>dall'antenna-Y di TDS. . . . .   | 8  |
| 9  | Moving Window Averaging con $N_l = N_r = 25$ sulla stessa IES della figura 8. . . . .   | 8  |
| 10 | spettro del pacchetto d'onda di Langmuir di figura 3 eseguito con Welch Data Win-<br>dowing (in alto) e senza Data Windowing (in basso). . . . .  | 11 |
| 11 | spettro del pacchetto d'onda di Langmuir di figura 16 eseguito con overlap (in alto) e<br>senza overlap (in basso). . . . .   | 12 |
| 12 | IES del 96/02/02 alle 12:04:36.554 UT (a sinistra), e spettro associato (a destra). . . . .   | 13 |
| 13 | pacchetto d'onda di Langmuir del 95/05/24 alle 17:53:57.630 UT (a sinistra), e spettro<br>associato (a destra). . . . .   | 13 |
| 14 | derivata (a sinistra) di una IES del 96/02/02 alle 13:13:43.546 UT (a destra): in rosso<br>eseguita da IDL, in blu da noi con lo schema a nove punti della tab. 2. . . . .  | 17 |
| 15 | derivata (a sinistra) di una IES del 97/09/16 alle 11:23:35.996 UT (a destra): in rosso<br>eseguita da IDL, in blu da noi con lo schema a nove punti della tab. 2. . . . .  | 17 |
| 16 | Pacchetto di Langmuir del 96/02/02 alle 15:50:13.660 UT (a sinistra), e suo integrale<br>(a destra): è ben visibile una pendenza media costante nella serie integrale. . . . .  | 18 |
| 17 | Come per la serie di fig. 16, ma per l'IES di fig. 3. . . . .   | 18 |
| 18 | autocorrelazione (a sinistra) dell'IES della fig. 3 (a destra). . . . .   | 19 |
| 19 | autocorrelazione (a sinistra) del pacchetto di Langmuir della fig. 13 (a destra). . . . .   | 20 |
| 20 | autocorrelazione (a sinistra) dell'IES della fig. 12 (a destra). . . . .  | 20 |
| 21 | autocorrelazione (a sinistra) del pacchetto di Langmuir della fig. 16 (a destra). . . . .   | 20 |
| 22 | esponenti di Lyapunov (con errori) del pacchetto di Langmuir di fig 16. Con linea<br>tratteggiata sono riportati anche gli esponenti calcolati per la serie temporale invertita:<br>come si vede un esponente è spurio, e gli altri sono positivi, quindi la serie è caotica. . . . . | 24 |

|    |   |    |
|----|---|----|
| 23 | esponenti di Lyapunov (ed errori) di un altro pacchetto di Langmuir del 96/02/02 alle 22:37:41.881 UT. Con linea tratteggiata sono riportati anche gli esponenti calcolati per la serie temporale invertita: anche questa ne ha uno spurio, e due positivi, quindi è caotica. . . . .       | 25 |
| 24 | esponenti di Lyapunov (in alto) di una IES del 96/02/02 alle 11:38:30.714 UT (in basso). Con linea tratteggiata sono riportati anche gli esponenti calcolati per la serie temporale invertita. Come si osserva gli esponenti sono tutti puri e positivi, quindi la serie è caotica. . . . . | 26 |

## Elenco delle tabelle

|   |  |    |
|---|--|----|
| 1 | Esempio di matrice $\mathbf{A}$ con $N_r = N_l = 1$ . . . . .                        | 16 |
| 2 | tabella dei coefficienti che abbiamo utilizzato: $N_r = 5, N_l = 3, n = 9$ . . . . . | 16 |