Multidimensional D^2 phase dispersion statistic

Nigul Olspert

AstroMHD Research Group Meeting

Monday 22nd August, 2016

Harmonics and periods

Spectral analysis

Thinking in terms of harmonics

- Power spectral density $\rightarrow$ periodogram
- Unevenly sampled data: Lomb-Scargle periodogram

PDM methods

Thinking in terms of periods

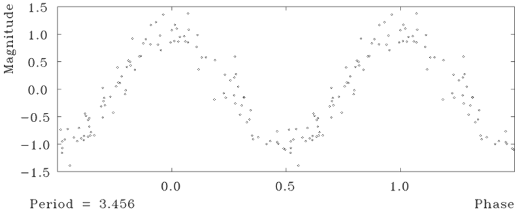
- By default support uneven sampling
- Generalizable to multiple dimensions

$$SS(P) = \sum_{i=1}^{n} (y_i(P) - \hat{y}_i(P))^2 \tag{1}$$

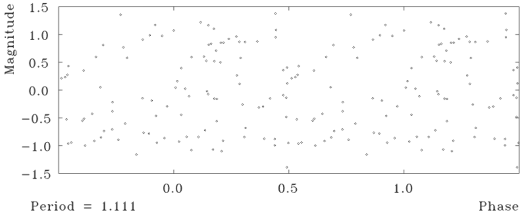
Lafler & Kinman: $\hat{y}_i(P) = y_{i-1}(P)$, Renson: weighted version
Stellingwerf: $\hat{y}_i(P) = \overline{y}_i(P)$, where $\overline{y}_i(P)$ is a bin mean.

Examples of phase diagram

Correct period



Wrong period



From periods to cycles

D^2 method

Thinking in terms of cycles

$$D^2(P, t_{\mathrm{coh}}) = \frac{\sum\limits_{i,j>i} g(t_i, t_j, P, t_{\mathrm{coh}}) ||\boldsymbol{f}(t_i) - \boldsymbol{f}(t_j)||^2}{2\sigma^2 \sum\limits_{i,j>i} g(t_i, t_j, P, t_{\mathrm{coh}})}, \qquad (2)$$

where $g = g_1(t_i, t_j, P) g_2(t_i, t_j, t_{\mathrm{coh}})$

- Essentially a generalized PDM method
- In addition to selection function g_1 in phase, introduces a selection function g_2 in time lag

We use:

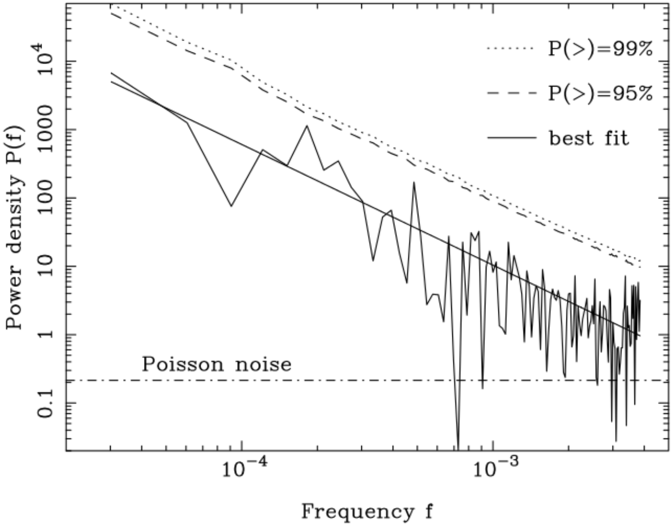
$g_1 = 2\cos(2\pi\nu(t_j - t_i))) + 1$

$g_2 = \exp(-\ln 2((t_j - t_i)/t_{\mathrm{coh}})^2)$

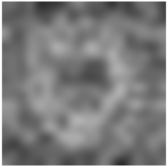
Significance estimation

- Null hypothesis: Gaussian white noise
- Stellingwerf showed that the spectral line has then *F* distribution.
- In many cases analytical form is hard or impossible to derive $\rightarrow$ use permutation test
- Is white noise correct null hypothesis?
- Also in red noise (or Brownian noise) patterns of periodicity can occur.
- Red noise is a special case of AR(1) process $X_t = c + \varphi X_{t-1} + \epsilon_t$ with $\varphi > 0$
- How to use red noise as null hypothesis?
- What if we could resample from the same process instead?

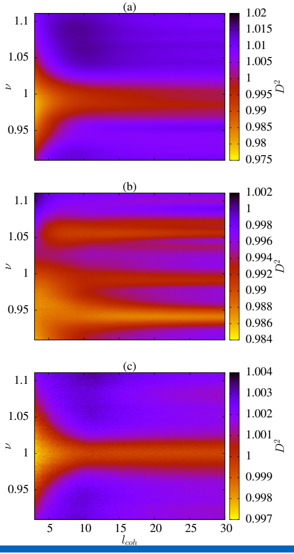
Example of red noise spectrum



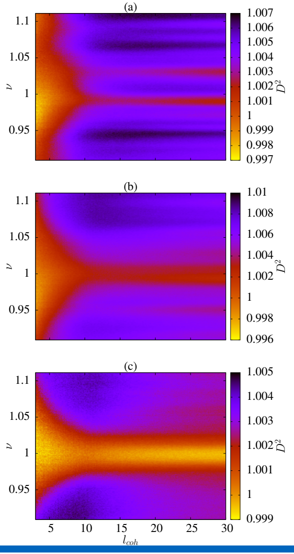
Test cases



Rotating particle

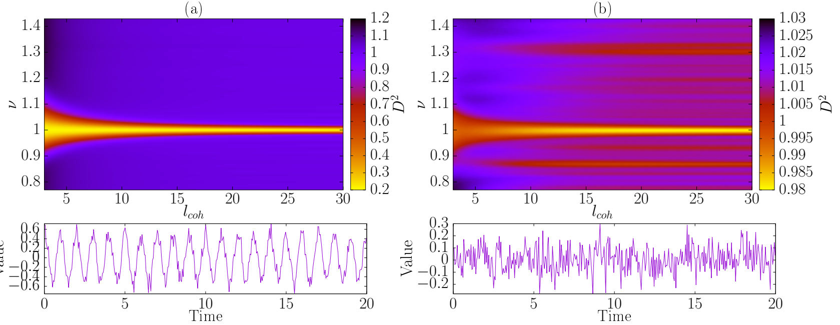


Oscillations in a box



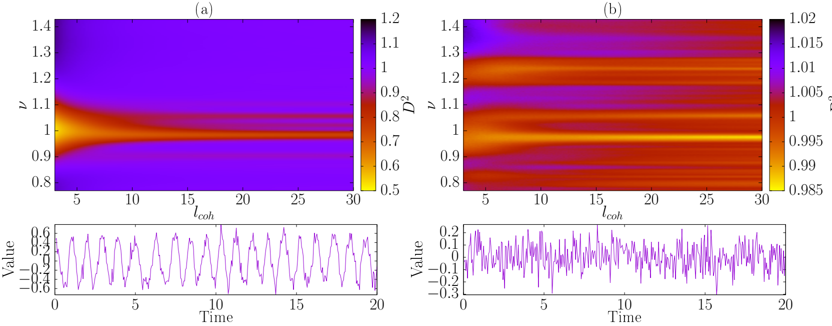
Patterns

Periodic signal



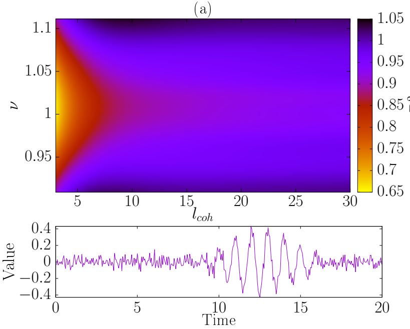
Patterns

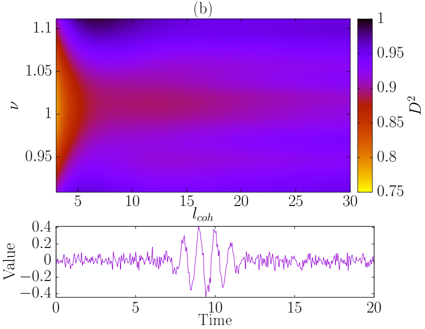
Cyclic signal



Patterns

Temporary signal





Cycles in PENCIL-Millennium data

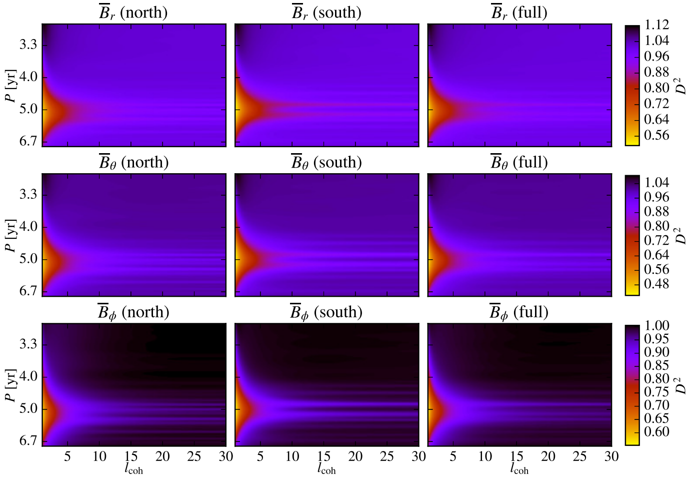
Mean cycle length estimates

Cycle no	B_r		B_θ		B_ϕ	
	N	S	N	S	N	S
I	0.47	0.47	0.48	0.48	0.46	0.46
II	5.12	4.98	5.13	4.98	5.17	5.02
III	*49.2*	*43.0*	46.2	40.2	50.8	46.06
IV	*108.4*	*105.1*	*108.0*	*106.0*	*107.5*	*104.1*

Notes: the numbers in *italic* represent cycles appearing only in the bottom of the convection zone, otherwise the cycle exists in the full hemisphere

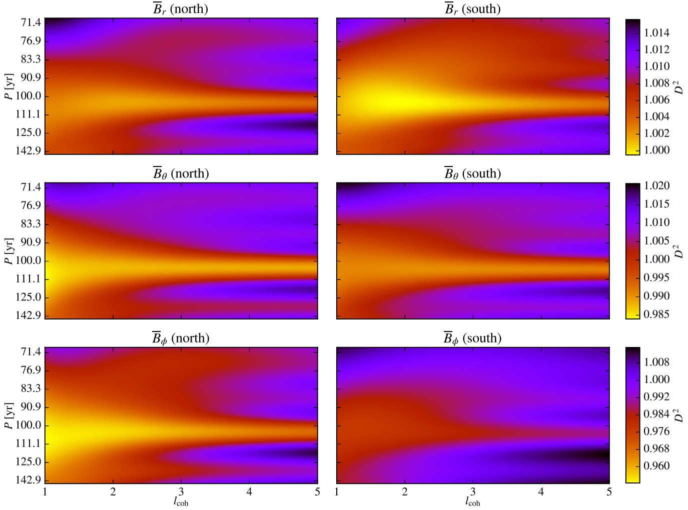
Cycles in PENCIL-Millennium data

5 year cycle



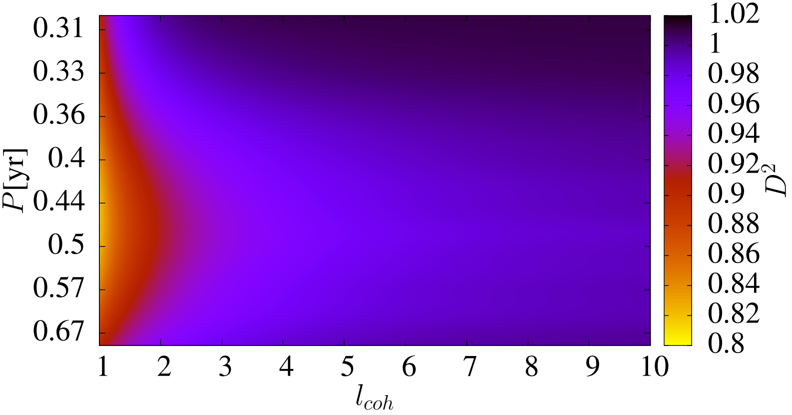
Cycles in PENCIL-Millennium data

100 year cycle



Cycles in PENCIL-Millennium data

0.5 year cycle



Cycles in PENCIL-Millennium data

50 year cycle