Lecciones en Astroinformática Avanzada (Semester 1 2024)

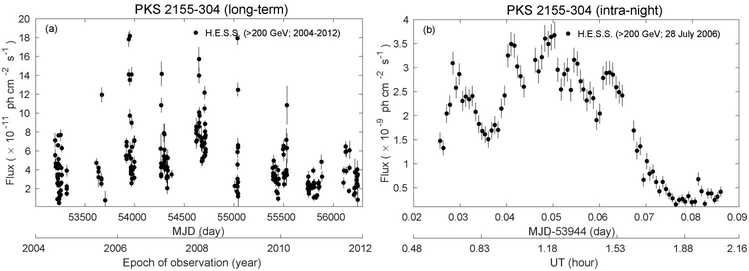
Stochastic Processes in Time Series

Nina Hernitschek
Centro de Astronomía CITEVA
Universidad de Antofagasta

June 18, 2024

Stochastic Processes

Despite light curves often show a characteristic behavior for a certain type of variable source, there are cases in which our best description of a light curve is that of a **stochastic process**.



Light Curve of the PKS 2155-304 hyper-high energy blazar

To successfully work with astronomical light curves from such as quasars, we have to solve this problem of fitting stochastic models to these data.

recap: Time Series

A time series is a sequence of random variables $\{\mathbf{X}_t\}_{t=1,2,\cdots}$.

Thus, a time series is a **series of data points ordered in time**. The time of observations provides a source of additional information to be analyzed.

recap: Time Series

A time series is a sequence of random variables $\{\mathbf{X}_t\}_{t=1,2,\cdots}$.

Thus, a time series is a **series of data points ordered in time**. The time of observations provides a source of additional information to be analyzed.

Since there may be an infinite number of random variables, we consider **multivariate distributions of random vectors**, that is, of finite subsets of the sequence $\{\mathbf{X}_t\}_{t=1,2,\cdots}$.

recap: Time Series

A time series is a sequence of random variables $\{\mathbf{X}_t\}_{t=1,2,\cdots}$.

Thus, a time series is a **series of data points ordered in time**. The time of observations provides a source of additional information to be analyzed.

Since there may be an infinite number of random variables, we consider **multivariate distributions of random vectors**, that is, of finite subsets of the sequence $\{\mathbf{X}_t\}_{t=1,2,\cdots}$.

A **time series model** for the observed data $\{x_t\}$ is defined to be a specification of all of the joint distributions of the random vectors $\mathbf{X} = (X_1, \cdots, X_n)^{\mathrm{T}}$, $n = 1, 2, \cdots$ of which $\{x_t\}$ are possible realizations, that is, at all of these probabilities

$$P(X_1 \leq x_1, \cdots, X_n \leq x_n), \; -\infty < x_1, \cdots, x_n < \infty, \\ n = 1, 2, \cdots .$$

Covariance Matrix

For a N-dimensional random vector $\mathbf{X}$ (such as a time series), one can calculate the **covariance matrix** which gives the covariance between each pair of elements of a given random vector:

If the entries in the column vector $\mathbf{X} = (X_1, X_2, ..., X_n)^{\mathrm{T}}$
are random variables, each with finite variance and expected value, then the covariance matrix $C_{\mathbf{XX}}$ is the matrix whose (i, j) entry is the covariance

$$C_{X_i X_j} = \mathrm{cov}[X_i, X_j] = \mathrm{E}[(X_i - \mathrm{E}[X_i])(X_j - \mathrm{E}[X_j])]$$

where the operator E denotes the **expectation value** of its argument.

The diagonal of this matrix gives the **variance** of the x_i.

$$C_{\mathbf{XX}} = \begin{bmatrix} Var(x_1) & \dots & Cov(x_1, x_n) \\ \vdots & \ddots & \vdots \\ Cov(x_n, x_1) & \dots & Var(x_n) \end{bmatrix}$$

Cross-Covariance Matrix

By comparison, the cross-covariance matrix between two vectors $\mathbf{X}, \mathbf{Y}$ is

$$\operatorname{cov}(\mathbf{X}, \mathbf{Y}) = \mathrm{C}_{\mathbf{XY}} = E\left[(\mathbf{X} - E[\mathbf{X}])(\mathbf{Y} - E[\mathbf{Y}])^{\mathrm{T}}\right].$$

In detail, the cross-covariance matrix is calculated as:

variance:

$$\operatorname{var}(\mathbf{X}) = \frac{\sum_i^N (X_i - \bar{X})^2}{N-1}$$

covariance:

$$\operatorname{cov}(\mathbf{X}, \mathbf{Y}) = \frac{\sum_i^N (X_i - \bar{X}) \cdot (Y_i - \bar{Y})}{N-1}$$

cross-covariance matrix:

$$C(\mathbf{X}, \mathbf{Y}) = \begin{bmatrix} var(\mathbf{X}) & cov(\mathbf{X}, \mathbf{Y}) \\ cov(\mathbf{X}, \mathbf{Y}) & var(\mathbf{Y})) \end{bmatrix}$$

$$C(\mathbf{X}, \mathbf{Y}, \mathbf{Z}) = \begin{bmatrix} var(\mathbf{X}) & cov(\mathbf{X}, \mathbf{Y}) & cov(\mathbf{X}, \mathbf{Z}) \\ cov(\mathbf{X}, \mathbf{Y}) & var(\mathbf{Y}) & cov(\mathbf{Y}, \mathbf{Z}) \\ cov(\mathbf{X}, \mathbf{Z}) & cov(\mathbf{Y}, \mathbf{Z}z) & var(\mathbf{Z}) \end{bmatrix}$$

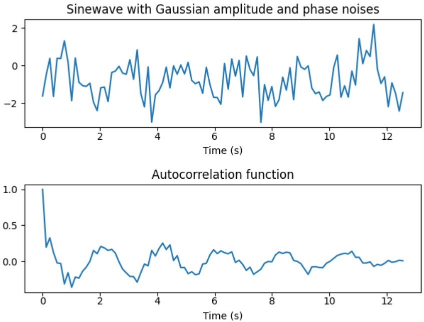
Autocovariance Function

Autocovariance is defined as the covariance between the value x_t and a value at a given time lag, $x_{t+\tau}$. It is usually denoted by the Greek letter γ:

$$\gamma(x_{t+\tau}, x_t) = \mathrm{E}\left[(X_{t+\tau} - \mu)\overline{(X_t - \mu)}\right]$$

Similar is the **autocorrelation**:

$$\mathrm{ACF}_{XX}(\tau) = \mathrm{E}\left[X_{t+\tau}\overline{X}_t\right]$$



Stationary Time Series

Definition

A time series $\{X_t\}$ is called **weakly stationary** or just stationary if

- $\mathrm{E}(X_t) = \mu_{X_t} = \mu < \infty$, that is, the expectation of X_t is finite and is not depending on t and
- $\gamma(x_{t+\tau}, x_t) = \gamma_\tau$, that is, for each τ, the autocovariance of $X_{t+\tau,X_t}$ is not depending on t.

Stationary Time Series

Definition

A time series $\{X_t\}$ is called **weakly stationary** or just stationary if

- $\mathrm{E}(X_t) = \mu_{X_t} = \mu < \infty$, that is, the expectation of X_t is finite and is not depending on t and
- $\gamma(x_{t+\tau}, x_t) = \gamma_\tau$, that is, for each τ, the autocovariance of $X_{t+\tau,X_t}$ is not depending on t.

A more restrictive definition of stationarity involves all the multivariate distributions of the subsets of time series random variables.

Definition

A time series $\{X_t\}$ is called **strictly stationary** if the random vectors $(X_{t_1}, \cdots, X_{t_n})^{\mathrm{T}}$ and $(X_{t_1+\tau}, \cdots, X_{t_n+\tau})^{\mathrm{T}}$ have the same joint distribution for all sets of indices $\{t_1, \cdots t_n\}$ and for all integers τ and $n > 0$. It is written as

$$(X_{t_1}, \cdots, X_{t_n})^{\mathrm{T}} \stackrel{\mathrm{d}}{=} (X_{t_1+\tau}, \cdots, X_{t_n+\tau})^{\mathrm{T}},$$

where $\stackrel{\mathrm{d}}{=}$ means *equal in distribution*.

Stationary Time Series

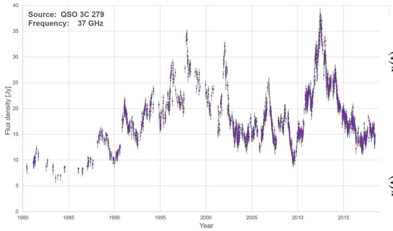
Properties of a Strictly Stationary Time Series:

- The random variables X_t are identically distributed for all t.
- Pairs of random variables $(X_t, X_{t+\tau})^{\mathrm{T}}$ are identically distributed for all t and τ, that is $(X_t, X_{t+\tau})^{\mathrm{T}} \stackrel{d}{=} (X_1, X_{1+\tau})^{\mathrm{T}}$
- The series X_t is a weakly stationary time series if $\mathrm{E}(X_t^2) < \infty$ for all t.
- Weak stationarity does not imply strict stationarity.

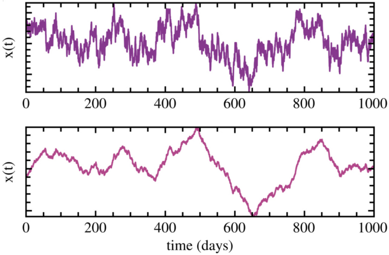
Describing Light Curves as Stochastic Processes

If a system is always variable, but the variability is not (infinitely) predictable, then we have a **stochastic process** (leading to a time series). These processes can also be characterized, but only statistically, not deterministically.

observed quasar light curve:



simulated light curves time series generated by a stochastic process (Moreno et al. 2019):



Stochastic Processes

Although the definition of a stochastic process varies, it is typically characterized as a collection of random variables indexed by some set.

Stochastic Processes

Although the definition of a stochastic process varies, it is typically characterized as a collection of random variables indexed by some set.

Stochastic variability includes behavior that is not predictable forever as in the periodic case, but unlike temporally localized events, variability is always present.

Stochastic Processes

Although the definition of a stochastic process varies, it is typically characterized as a collection of random variables indexed by some set.

Stochastic variability includes behavior that is not predictable forever as in the periodic case, but unlike temporally localized events, variability is always present.

Typically, the **underlying physics** is so complex that we cannot deterministically predict future values. Despite their seemingly irregular behavior, there are a number of ways on how to **quantify and characterize** the data.

Autoregressive Models

Processes that are not periodic, but that nevertheless *retain memory* of previous states, can be described in terms of **autoregressive models**.

Autoregressive models (AR) with dependencies on k past values are called **autoregressive process of order** k and denoted as AR(k). A generalization is called the continuous autoregressive process, CAR(k).

Autoregressive Models

For **linear regression**, we are predicting the dependent variable from the independent variable

$$y = mx + b.$$

For **auto-regression**, the dependent and independent variable is the same and we are predicting a future value of y based on k past values of y:

$$y_i = a_i y_{i-1} + \ldots = \sum_{j=1}^{k} a_j y_{i-j} + \epsilon_i$$

where a_i is the **lag coefficient**.

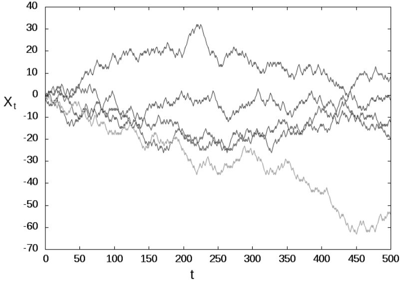
Autoregressive Models

A **random walk** is an example of such a process (with $a_i = 1$, $k = 1$); every value is given by the preceeding value plus noise.

The model can then be written as

$$X_t = X_{t-1} + Z_t,$$

where Z_t is a white noise variable with zero mean and variance σ^2. This model is *not stationary*.



different realizations of a 1D Random Walk time series with 500 time steps

Random Walk

Repeatedly substituting for past variables results in

$$
\begin{aligned}
X_t &= X_{t-1} + Z_t \\
&= X_{t-2} + Z_{t-1} + Z_t \\
&= X_{t-3} + Z_{t-2} + Z_{t-1} + Z_t \\
&\vdots \\
&= X_0 + \sum_{j=0}^{t-1} Z_{t-j}.
\end{aligned}
$$

If the initial value X_0 is fixed, then the mean value of X_t is equal to X_0, that is,

$$\mathrm{E}(X_t) = \mathrm{E}\left(X_0 + \sum_{j=0}^{t-1} Z_{t-j}\right) = X_0.$$

Random Walk

The mean is constant, but the variance and covariance depend both on time, not just on the lag.
Since the white noise variables Z_t are uncorrelated, we obtain

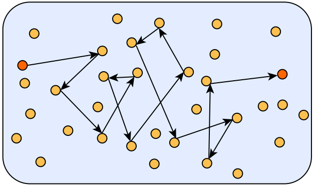
$$\begin{aligned}\mathrm{Var}(X_t) &= \mathrm{Var}\left(X_0 + \sum_{j=0}^{t-1} Z_{t-j}\right) = \mathrm{Var}\left(\sum_{j=0}^{t-1} Z_{t-j}\right) \\ &= \sum_{j=0}^{t-1} \mathrm{Var}\left(Z_{t-j}\right) = t\sigma^2\end{aligned}$$

and

$$\begin{aligned}\mathrm{Cov}(X_t, X_{t-\tau}) &= \mathrm{Cov}\left(\sum_{j=0}^{t-1} Z_{t-j}, \sum_{k=0}^{t-\tau-1} Z_{t-\tau-k}\right) \\ &= \mathrm{E}\left\{\left(\sum_{j=0}^{t-1} Z_{t-j}\right)\left(\sum_{k=0}^{t-\tau-1} Z_{t-\tau-k}\right)\right\} \\ &= \min(t, t-\tau)\sigma^2.\end{aligned}$$

Damped Random Walk

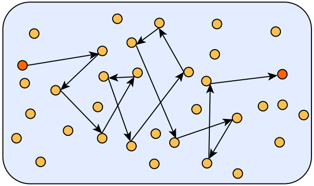
The **Ornstein-Uhlenbeck process** is a stochastic process that originally describes the velocity of a massive Brownian particle under the influence of friction.



Brownian motion is the random motion of particles suspended in a medium.

Damped Random Walk

The **Ornstein-Uhlenbeck process** is a stochastic process that originally describes the velocity of a massive Brownian particle under the influence of friction.



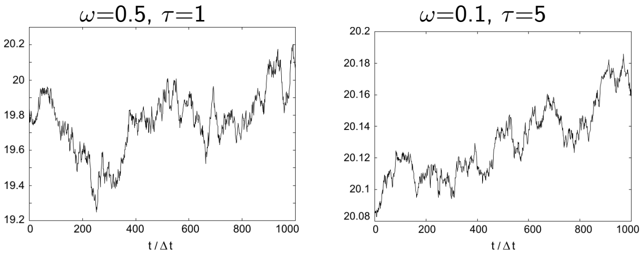
Brownian motion is the random motion of particles suspended in a medium.

The process can be considered as a modification of the random walk but with a **tendency to move back towards a central location**, with a greater **attraction** when the process is further away from the center. The process is stationary.

Damped Random Walk

An Ornstein-Uhlenbeck process x_t satisfies the following stochastic differential equation: $\mathrm{d}x_t = \tau(\mu - x_t)\mathrm{d}t + \omega\,\mathrm{d}W_t$

where $\tau > 0$, μ and $\omega > 0$ are parameters and W_t denotes the Wiener process (random walk in continuous time).
The stationary (long-term) variance is given by $\mathrm{Var}(x_t) = \frac{\omega^2}{2\tau}$.



two realizations of a 1D Damped Random Walk, each with a mean of 20

Damped Random Walk

The damped random walk can also be described by its covariance matrix:

$$S_{ij} = \sigma^2 \exp(-|t_{ij}/\tau|)$$

where σ and τ are the model parameters.
σ^2 controls the short timescale covariance ($t_{ij} \ll \tau$), which decays exponentially on a timescale given by τ which is called the characteristic timescale (relaxation time, or damping timescale).

With this, the **autocorrelation function for a damped random walk** is

$$\mathrm{ACF}_{\mathrm{DRW}}(t) = \exp(-t/\tau).$$

Moving Average Process

A **moving average (MA) process** is similar to an AR process, but the value at each time step depends not on the value of previous time step, but rather the **perturbations from previous time steps**.
MA processes are defined by

$$y_i = \epsilon_i + \sum_{j=1}^{q} b_j \epsilon_{i-j}.$$

So, for example, an MA(q=1) process would look like

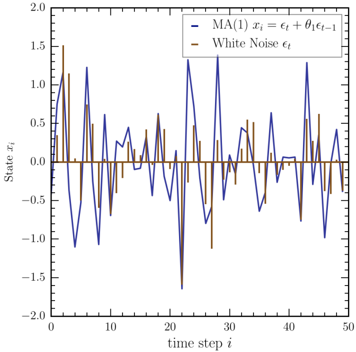
$$y_i = \epsilon_i + b_1 \epsilon_{i-1},$$

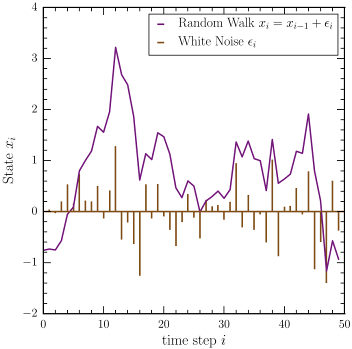
whereas an AR(p=2) process would look like

$$y_i = a_1 y_{i-1} + a_2 y_{i-2} + \epsilon_i$$

Moving Average Process

Thus in an MA process a 'shock/impulse' affects only the current value and q values into the future. In an *AR* process a 'shock/impulse' affects all future values. These two plots show the difference between an MA(1) process and an AR(1) (random walk) process:





credit: Moreno et al. (2019)

(Auto-)Correlation Function

problem:
We observe a (stochastically varying) quasar which has both **line and continuum emission**; the line emission is stimulated by the continuum.

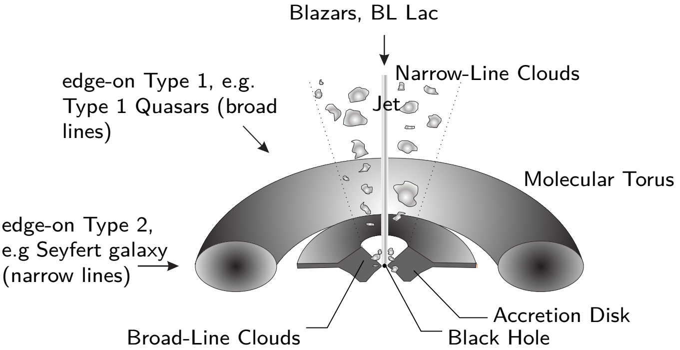
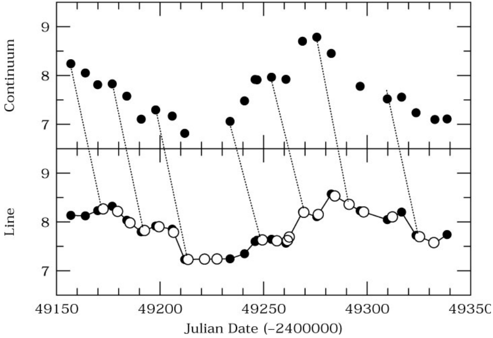


Illustration of the unified AGN model for the radio-quiet case. The classes of AGN are indicated by the viewing angle shown by arrows.

(Auto-)Correlation Function

The physical separation between the regions that produce each type of emission causes a delay between the light curves:



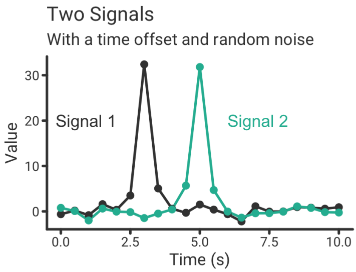
credit: Peterson+2001

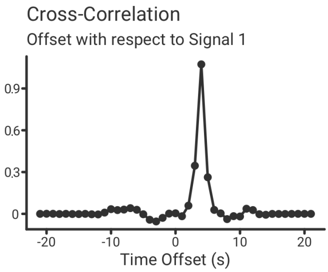
We measure line and continuum emission and want to compute the time lag t_{lag}.

Correlation Function

To find a time lag between two time series, we can compute the **correlation function**. If one time series is derived from another simply by shifting the time axis by t_{lag}, then their (cross-)correlation function will have a peak at $\Delta t = t_{\mathrm{lag}}$.

Computing the correlation function is basically the mathematical processes of convolution, i.e., sliding the two curves over each other and computing the degree of similarity for each step in time:





Correlation Function

The **(cross-)correlation function** between time series $f(t)$ and $g(t)$ is defined as

$$\mathrm{CCF}(\Delta t) = \frac{\lim_{T\to\infty} \frac{1}{T} \int_T f(t)\, g(t+\Delta t) dt}{\sigma_f \sigma_g}$$

σ_f and σ_g are the standard deviations of $f(t)$ and $g(t)$, respectively. With this normalization, the correlation function is unity for $\Delta t = 0$.

Correlation Function

The **(cross-)correlation function** between time series $f(t)$ and $g(t)$ is defined as

$$\mathrm{CCF}(\Delta t) = \frac{\lim_{T\to\infty} \frac{1}{T} \int_T f(t)\, g(t+\Delta t) dt}{\sigma_f \sigma_g}$$

σ_f and σ_g are the standard deviations of $f(t)$ and $g(t)$, respectively. With this normalization, the correlation function is unity for $\Delta t = 0$.

For the **autocorrelation function (ACF)**, we take our correlation function from above and set $f(t) = g(t)$.

Correlation Function

The **(cross-)correlation function** between time series $f(t)$ and $g(t)$ is defined as

$$\mathrm{CCF}(\Delta t) = \frac{\lim_{T\to\infty} \frac{1}{T} \int_T f(t)\, g(t+\Delta t) dt}{\sigma_f \sigma_g}$$

σ_f and σ_g are the standard deviations of $f(t)$ and $g(t)$, respectively. With this normalization, the correlation function is unity for $\Delta t = 0$.

For the **autocorrelation function (ACF)**, we take our correlation function from above and set $f(t) = g(t)$.

Whereas the correlation function yields information about a possible time lag between $f(t)$ and $g(t)$, the autocorrelation function yields information about the variable timescales present in a process. When $f(t)$ values are uncorrelated (e.g., due to white noise without any signal), $\mathrm{ACF}(\Delta t) = 0$ except for ACF(0)=1. For processes that *retain memory* of previous states only for some characteristic time τ, the autocorrelation function vanishes for $\Delta t \ll \tau$.

Correlation Function

Let $x(t_i)$ and $y(t_i)$ be two **discrete, evenly sampled time series**.
In this case, the cross-correlation function as a function of time lag τ is

$$\mathrm{CCF}(\tau) = \frac{1}{N}\sum_{i=1}^{N}\frac{[x(t_i)-\bar{x}]\,[y(t_i-\tau)-\bar{y}]}{\sigma_x\sigma_y},$$

Correlation Function

Let $x(t_i)$ and $y(t_i)$ be two **discrete, evenly sampled time series**.
In this case, the cross-correlation function as a function of time lag τ is

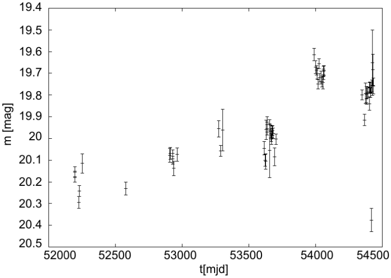
$$\mathrm{CCF}(\tau) = \frac{1}{N}\sum_{i=1}^{N}\frac{\left[x(t_i)-\bar{x}\right]\left[y(t_i-\tau)-\bar{y}\right]}{\sigma_x\sigma_y},$$

The **accuracy** with which the peak of a CCF can be determined when the data points are evenly spaced has been dealt with by Tonry et al. (1979). The accuracy which is achieved is proportional to the ratio of the height of the peak to the noise level in the CCF divided by the half-width at half-maximum of the peak in the CF.

Correlation Function for Unevenly Sampled Time Series

In practice, for time series being **long-time astronomical data** or other measurements, such as those produced by quasar monitoring programs, $x(t_i)$ and $y(t_j)$ are usually **not known at regular intervals of time**.

A considerable amount of theory has been developed for this case, how one can carries out time series analysis based on CCF even in this cases.



A quasar light curve from SDSS Stripe 82.

Correlation Function for Unevenly Sampled Time Series

There are in general two approaches for dealing with the uneven sampling: the **Discrete Cross Correlation Function (DCCF)** and interpolation methods like the **Interpolated Cross-Correlation Function (ICCF)**.

These two methods are briefly described in the following.

The Discrete Cross Correlation Function (DCCF)

In the DCCF method (Edelson+1988), a contribution to the DCCF is **calculated only using the actual data points**. Each pair of points, one from each of the light curves, gives one correlation value at a lag corresponding to the time separation. For two light curves with N and M data points respectively this gives an Unbinned Cross Correlation Function (UCCF),

$$\mathrm{UCCF}_{ij} = \frac{(x_i - \bar{x})(y_j - \bar{y})}{\sigma_x \sigma_y}.$$

The DCCF is then obtained by averaging the UCCF in time lag bins. This results in

$$\mathrm{UCCF}_{ij} = \frac{1}{N} \frac{(x_i - \bar{x})(y_j - \bar{y})}{\sigma_x \sigma_y}.$$

where N is the number of pairs.

A point to notice is that a large variation in observational coverage over the light curve can have a strong effect on the DCCF amplitudes.

The Interpolated Cross Correlation Function (ICCF)

In the ICCF method (Gaskell+1986), the light curve is **linearly interpolated** and resampled onto a regular grid. It is common to calculate the ICCF twice, where interpolation is done in each of the light curves, one at a time, and average these.
For two time series x_i and y_i, in the case of interpolation in time series y_i, the cross-correlation function is approximated by

$$\mathrm{ICCF}_{xy}(\Delta t) = \frac{1}{N} \sum_{i=1}^{N} \frac{(x_{t_i} - \bar{x})\{L\left[y_{t_i+\Delta t}\right] - \bar{y}\}}{\sqrt{\sum_{i=1}^{N}(x_{t_i} - \bar{x})^2}\sqrt{\sum_{i=1}^{N}(y_{t_i} - \bar{y})^2}}$$

where L indicates a piecewise linear interpolation of series y at time $t_i + \Delta t$ and N is the number of data points in x.
The piecewise linear interpolation of y is done by

$$L\left[y_{t_i+\Delta t}\right] = \begin{cases} y_{t_i+\Delta t} & \text{if observed for } t_i + \Delta t \\ y_{t_{i-1}} \frac{t_{i+1}-(t_i+\Delta t)}{t_{i+1}-t_{i-1}} + y_{t_{i+1}} \frac{(t_i+\Delta t)-t_{i-1}}{t_{i+1}-t_{i-1}} & \text{else} \end{cases}$$

where $y_{t_{i-1}}$ and $y_{t_{i+1}}$ are indicating the nearest data points available.

Structure Function

The **structure function** is another quantity that is frequently used in astronomy and is related to the ACF:

$$\mathrm{SF}(\Delta t) = \mathrm{SF}_\infty [1 - \mathrm{ACF}(\Delta t)]^{1/2},$$

where SF_∞ is the standard deviation of the time series as evaluated on timescales much larger than any charateristic timescale, τ.

The ACF for a Damped Random Walk (DRW) is given by

$$\mathrm{ACF}(t) = \exp(-t/\tau),$$

where τ is the characteristic timescale (i.e., the damping timescale). Remember that a DRW modeled as an AR(1) has $a_1 = \exp(-1/\tau)$. The **structure function for a DRW** can then be written as

$$\mathrm{SF}(t) = \mathrm{SF}_\infty [1 - \exp(-t/\tau)]^{1/2}.$$

The Problem

determination of Black Hole masses

needed for

- understanding the AGN phenomena
- cosmological evolution of BHs
- coevolution of AGNs and their host galaxies

techniques

- very nearby galaxies: stellar and gas dynamics near their center
- more distant objects: direct measurements are rare, mostly simulations, gravitational lensing and others
- special case of AGN: **reverberation mapping**

Reverberation Mapping

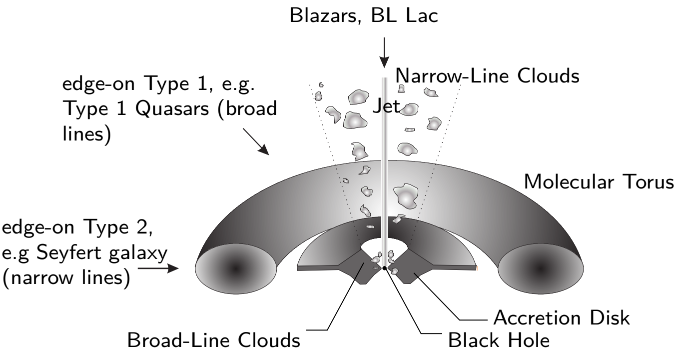
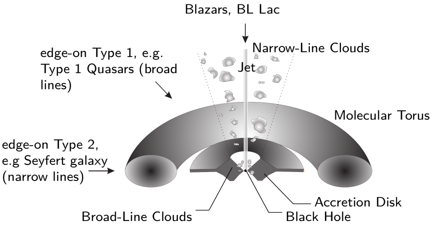


Illustration of the unified AGN model for the radio-quiet case. The classes of AGN are indicated by the viewing angle shown by arrows.

Reverberation Mapping



- continuum radiation from accretion disk excites cold Broad Line Region clouds ⇒ broad emission lines
- kinematics and geometry of BLR ⇒ changes in Broad Line Region excitation and luminosities
- finite light travel time ⇒ delay of Broad Line Region luminosity variations with respect to accretion disk luminosity variation

Reverberation Mapping

Broad Line Region (BLR) size: $R_{\rm BLR} = c\ \tau_{\rm delay}$

$\Rightarrow$ assuming Keplerian orbits for the BLR cloud:

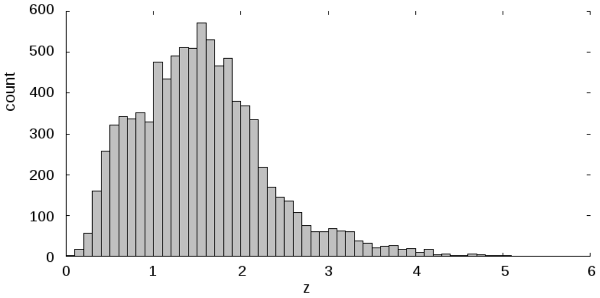
$$M_{\rm BH} = f\ \frac{\Delta V^2 c \tau_{\rm delay}}{G}$$

where f is a proportionality factor of order unity depending on BLR geometry and kinematics (Kaspi et al. 2000)

SDSS Stripe 82

SDSS Stripe 82 (S82) contains, among other sources, light curves for all spectroscopically confirmed QSOs within that region (22h 24m $< R.A. <$ 04h 08m and $|Dec| < 1.27$ deg, about 290 deg^2). The total number of QSOs is 9,258, and the observations are spaced out over $\sim$10 years in yearly "seasons" about 2-3 months long. The redshift is $\lesssim 5$.

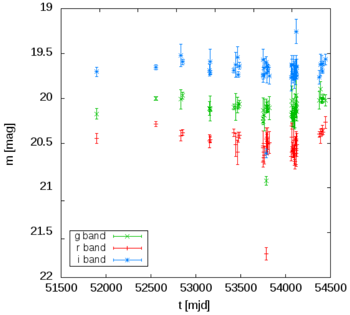
redshift distribution of 9,120 quasars of SDSS Stripe 82

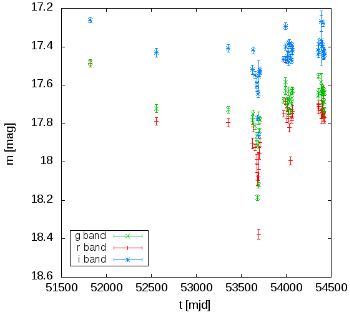


SDSS Stripe 82

- each $\sim$ 60 epochs mag measurements precise to 2 % over $\sim$ 8 years
- five bandpasses between 3,500 $\mathring{A}$ and 9,000 $\mathring{A}$
- in some bands, BLR flux contributes up to 20 %
- single-epoch spectra, for some objects multi-epoch
- broad band data for Hα, Hβ, Mg II
- cross-matched with spectra and Catalog of Quasar properties

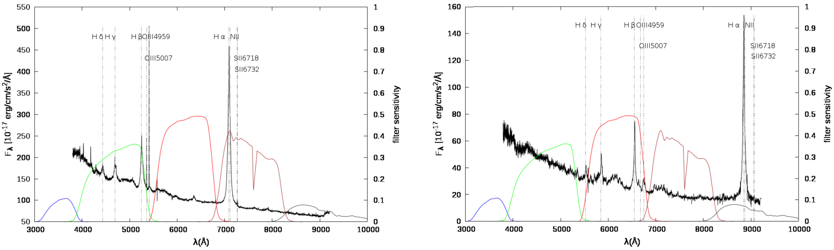
Example Light Curves





two example light curves from SDSS S82 quasars, only three bands are shown

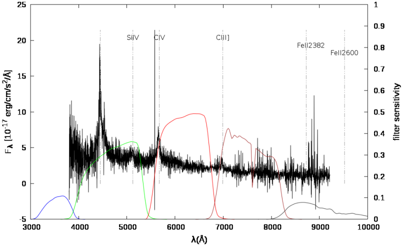
Example Spectra



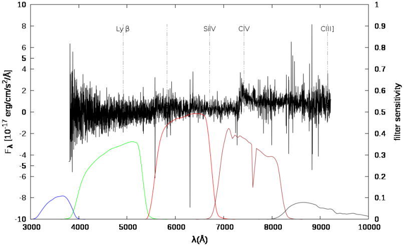
(a) (b)

two example spectra from SDSS S82 quasars, filter curves overplotted
(a) z=0.0802. g band: Hβ, Hγ, O III λ4959, r and z band: continuum
(b) z=0.347337. z band: Hα, r band: Hβ, but all in low transmission regions

Example Spectra



(c)



(c)

two example spectra from SDSS S82 quasars, filter curves overplotted

(c) z=2.6613. None of the available emission lines is in a useful filter band regime.

(d) z=3.7927. None of the available emission lines is in a useful filter band regime.

Methodology

approach assumes that all emission-line light curves are scaled, smoothed and displaces versions of the continuum

flux model:

$$
\begin{aligned}
f_x(t) &= f_x^c(t) && \text{continuum only band} \\
f_y(t) &= f_y^c(t) + f_y^e(t) && \text{continuum and emission line contribution band} \\
&= s \cdot f_x^c(t) + e \int \mathrm{d}\tau_{\mathrm{delay}} \Psi(\tau_{\mathrm{delay}}) f_x^c(t - \tau_{\mathrm{delay}})
\end{aligned}
$$

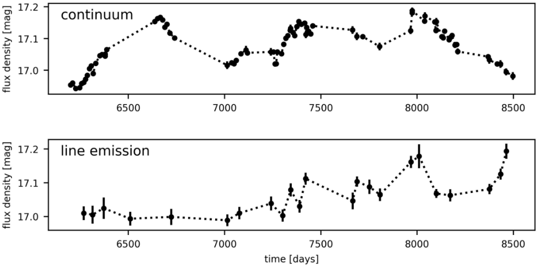
superscripts ^{c} and ^{e} denote continuum or emission line contributions, $\Psi(\tau_{\mathrm{delay}})$ is the transfer function and s, e are scaling factors

Classic Reverberation Mapping

The **classic approach** for reverberation mapping relies on the ACF and CCF.

The problem with this approach is: It cannot deal with irregular sampling of time series, like often present in astronomical surveys.

To cope with this problem, various approaches for interpolation were developed, e.g. ICCF (interpolated cross-correlation function). With a linear interpolation, they only work when the sampling is not too irregular and when the cadence is high.



Stochastical Reverberation Mapping

Taking into account the AGN continuum variability, which can be modelled as a stochastic process and specifically as a damped random walk, the line-emission light curve can be treated as a scaled, smoothed, and time-lagged response to the continuum emission.

The **stochastic approach** that is capable to:

- not only interpolate between data points, but also make self-consistently estimates and include these uncertainties in the interpolation
- handle transfer functions $\Psi(\tau_{\text{delay}})$ instead of simply a τ_{delay}
- separate light-curve means and systematic errors in flux calibration from variability signals and measurement noise in a self-consistent way
- derive simultaneously the lags of multiple emission lines
- provide statistical confidence limits on all estimated parameters

Stochastical Reverberation Mapping

Taking into account the AGN continuum variability, which can be modelled as a stochastic process and specifically as a damped random walk, the line-emission light curve can be treated as a scaled, smoothed, and time-lagged response to the continuum emission.

The **stochastic approach** that is capable to:

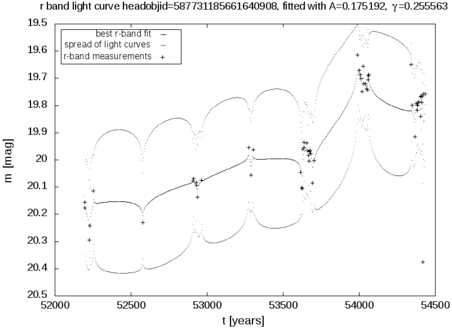
- not only interpolate between data points, but also make self-consistently estimates and include these uncertainties in the interpolation
- handle transfer functions $\Psi(\tau_{\mathrm{delay}})$ instead of simply a τ_{delay}
- separate light-curve means and systematic errors in flux calibration from variability signals and measurement noise in a self-consistent way
- derive simultaneously the lags of multiple emission lines
- provide statistical confidence limits on all estimated parameters

⇒ handle sparsely sampled data!

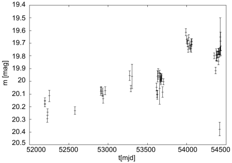
The General Approach

We fit a light curve by maximizing the likelihood that the light curve can be fitted by a model:

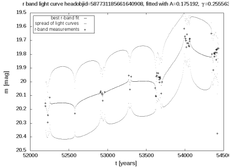
approach builds on a fitting method developed of Rybicki et al. (1994) and Zu et al. (2011), applied to broad band photometry with a lot of modifications (see Hernitschek et al. 2015)

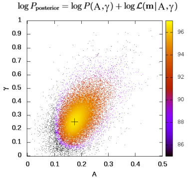


The General Approach

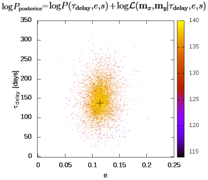


$\Downarrow$





$\Downarrow$



model continuum as a Gaussian stochastic process

$$\log P_{\text{post}} = \log P(\mathbf{p}) + \log \mathcal{L}(\mathbf{m}|\mathbf{p})$$

where **p** are the structure function parameters and **m** the measured light curve points

model band with emission line contribution as scaled version of pure continuum plus scaled, smoothed and displaced version of the continuum

$$\log P_{\text{post}} = \log P(\tau_{\text{delay}}, e, s) + \log \mathcal{L}(\mathbf{m}_x, \mathbf{m}_y|\tau_{\text{delay}}, e, s)$$

Modeling the Continuum

model the continuum-only band x

describe quasar continuum light curve as Gaussian stochastic process (e.g., Kozlowski et al. 2009, MacLeod et al. 2012

- damped random walk (Kelly et al. 2009)
- power-law structure function model (Schmidt et al. 2010)

$\Rightarrow$ continuum model is characterized by a variance matrix $C_{xx}^{cc}(\Delta t)$

Modeling the Emission Line

Model the band with continuum and emission line contribution y

$$f_y^e(t) = e \int \mathrm{d}\tau_{\mathrm{delay}} \Psi(\tau_{\mathrm{delay}}) f_x^c(t - \tau_{\mathrm{delay}})$$

The emission-line covariance matrix C^{ee} is then given by

$$\begin{aligned} C_{yy}^{ee}(\Delta t) &= \langle f_y^e(t), f_y^e(t + \Delta t)\rangle \\ &= e^2 \int \mathrm{d}\tau_{\mathrm{delay},1} \int \mathrm{d}\tau_{\mathrm{delay},2} \Psi(\tau_{\mathrm{delay},1}) \langle f_y^c(t\tau_{\mathrm{delay},1}) f_y^c(t + \Delta t - \tau_{\mathrm{delay},2})\rangle \end{aligned}$$

Modeling the Emission Line

continuum-emission cross terms are given by

$$C_{yy}^{ec/ce}(\Delta t) = e \int \mathrm{d}\tau_{\mathrm{delay}} \Psi(\tau_{\mathrm{delay}}) C_{yy}^{ee}(\Delta t \pm \tau_{\mathrm{delay}})$$
$$= s \cdot e \int \mathrm{d}\tau_{\mathrm{delay}} \Psi(\tau_{\mathrm{delay}}) C_{xx}^{cc}(\Delta t \pm \tau_{\mathrm{delay}})$$

$\Rightarrow$ covariance matrix for the x band continuum and y band continuum plus emission line fluxes

$$C = \begin{pmatrix} C_{xx}^{cc} & C_{xy}^{c,(e+c)} \\ C_{yx}^{(e+c),c} & C_{yy}^{(e+c),(e+c)} \end{pmatrix}$$

Maximum Likelihood Estimation

structure function parameter estimation: evaluate logarithmic posterior probability distribution

$$\log P_{\mathrm{posterior}} = \log P(\mathbf{p}) + \log \mathcal{L}(\mathbf{m}|\mathbf{p})$$

where $\mathbf{p}$: structure function parameters, $\mathbf{m}$: light curve measurements

τ_{delay} estimation: evaluate logarithmic posterior probability distribution

$$\log P_{\mathrm{posterior}} = \log P(\tau_{\mathrm{delay}}, e, s) + \log \mathcal{L}(\mathbf{m}_x, \mathbf{m}_y|\tau_{\mathrm{delay}}, e, s)$$

likelihood $\mathcal{L}$ from covariance matrix C and data $\mathbf{m}$ (Zu et al. 2011)

$$\mathcal{L} \equiv |S+N|^{-1/2}|L^T C^{-1} L|^{-1/2} \exp\left(-\frac{\mathbf{m}^T C_{\perp}^{-1}\mathbf{m}}{2}\right)$$

Selecting Data

this is applied to light curves being selected by

- redshift

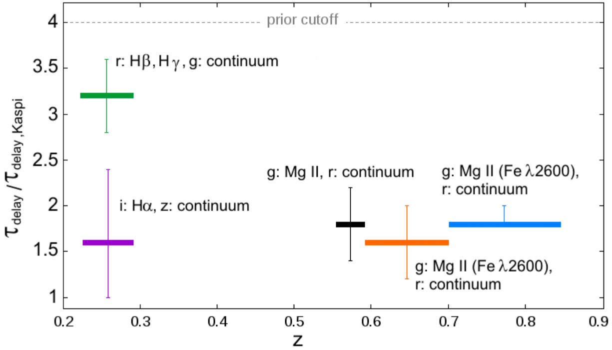
  one band emission-line free, one with emission line contribution

- time sampling

  has great influence on the possibility of estimating τ_{delay}

$$\Longrightarrow \quad \begin{array}{l} 71 \text{ light curves in } z = 0.555 - 0.591 \\ 37 \text{ light curves in } z = 0.225 - 0.291 \\ 21 \text{ light curves in } z \lesssim 0.2 \end{array}$$

Results



1 spectrum and $\sim$ 60 epochs each yield

- marginal τ_{delay} estimates with mock data & real data
- solid τ_{delay} estimate in ensemble average
- application to 323 quasars in SDSS S82
- future: more spectroscopic epochs

Summary

Stochastic processes are a method to fit time series that show an irregular behavior that is otherwise not possible to describe.

One application is reverberation mapping, in which quasar light curves in different wavebands need to be compared.

Other applications lie in finding a general description for light curves with a model that can be fit to any kind of time series. This can be used as a first step for extracting features to classify light curves in large surveys, before more specific feature extraction methods are applied to subsets of the survey data.