Microlensing: Theory, Practice, Results, Future

Lecture 4

Nicholas James Rattenbury

JODRELL BANK CENTRE FOR ASTROPHYSICS
THE UNIVERSITY OF MANCHESTER

Outline of Lectures

Lecture 1: Microlensing History and Theory

Outline of Lectures

Lecture 1: Microlensing History and Theory

- Background
- Motivation/Goals
- Early results
- Evolution of a field
- Basic microlensing theory

Outline of Lectures

Lecture 1: Microlensing History and Theory

Lecture 2: Beyond the Single lens

Outline of Lectures

Lecture 1: Microlensing History and Theory

Lecture 2: Beyond the Single lens

- Finite source star
- Limb Darkening
- Blending
- Parallax
- Xallarap

Outline of Lectures

Lecture 1: Microlensing History and Theory

Lecture 2: Beyond the Single lens

Lecture 3: Planetary Microlensing - I

Outline of Lectures

Lecture 1: Microlensing History and Theory

Lecture 2: Beyond the Single lens

Lecture 3: Planetary Microlensing - I

- Binary lens microlensing
- Extreme mass ratio microlensing
- Theoretical tools of trade: caustics
- Planetary microlensing regiemes
- General rules

Outline of Lectures

Lecture 1: Microlensing History and Theory

Lecture 2: Beyond the Single lens

Lecture 3: Planetary Microlensing - I

Lecture 4: Planetary Microlensing - II

Outline of Lectures

Lecture 1: Microlensing History and Theory

Lecture 2: Beyond the Single lens

Lecture 3: Planetary Microlensing - I

Lecture 4: Planetary Microlensing - II

- Capabilities, detection limits
- Detection
- Modelling

Detection techniques

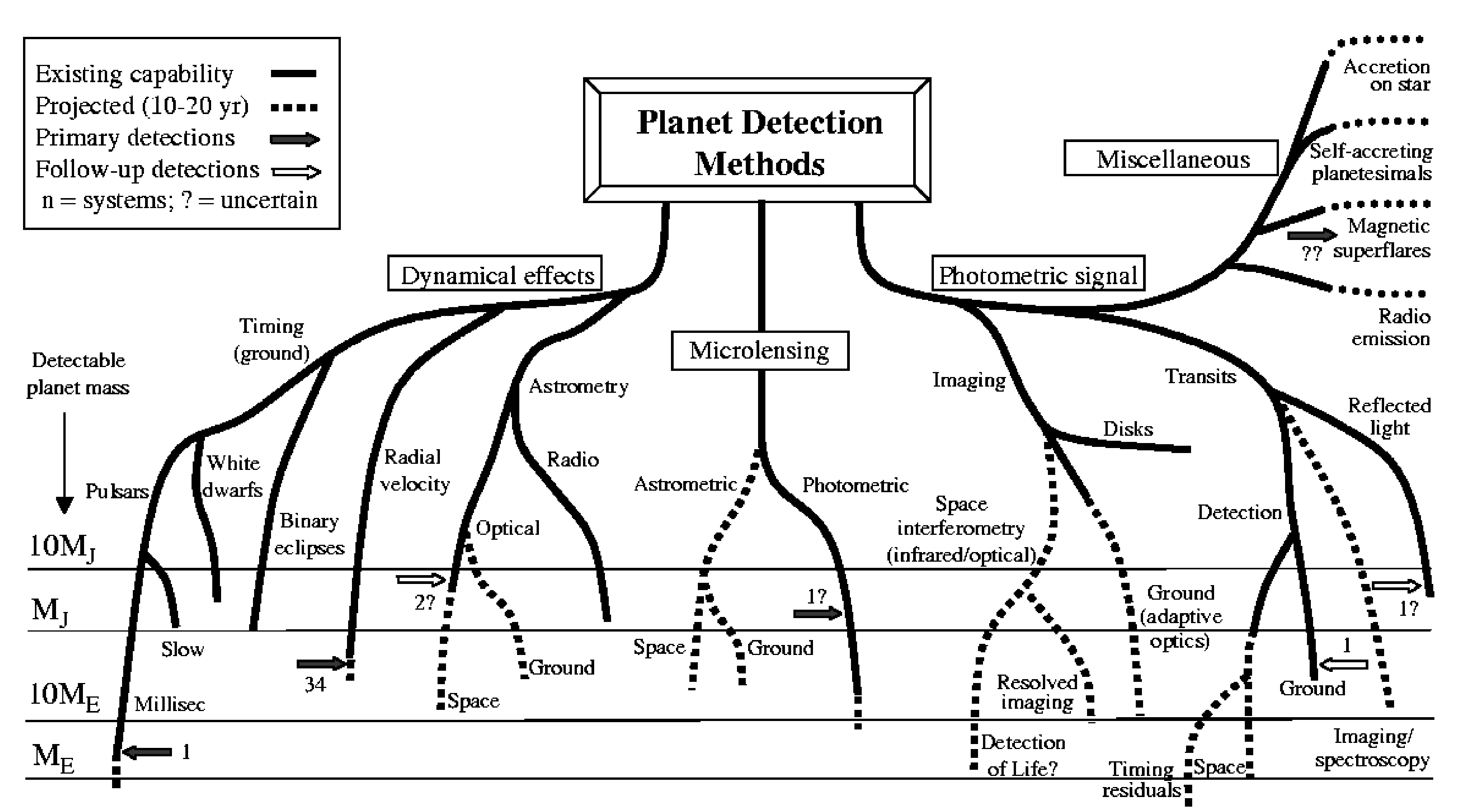
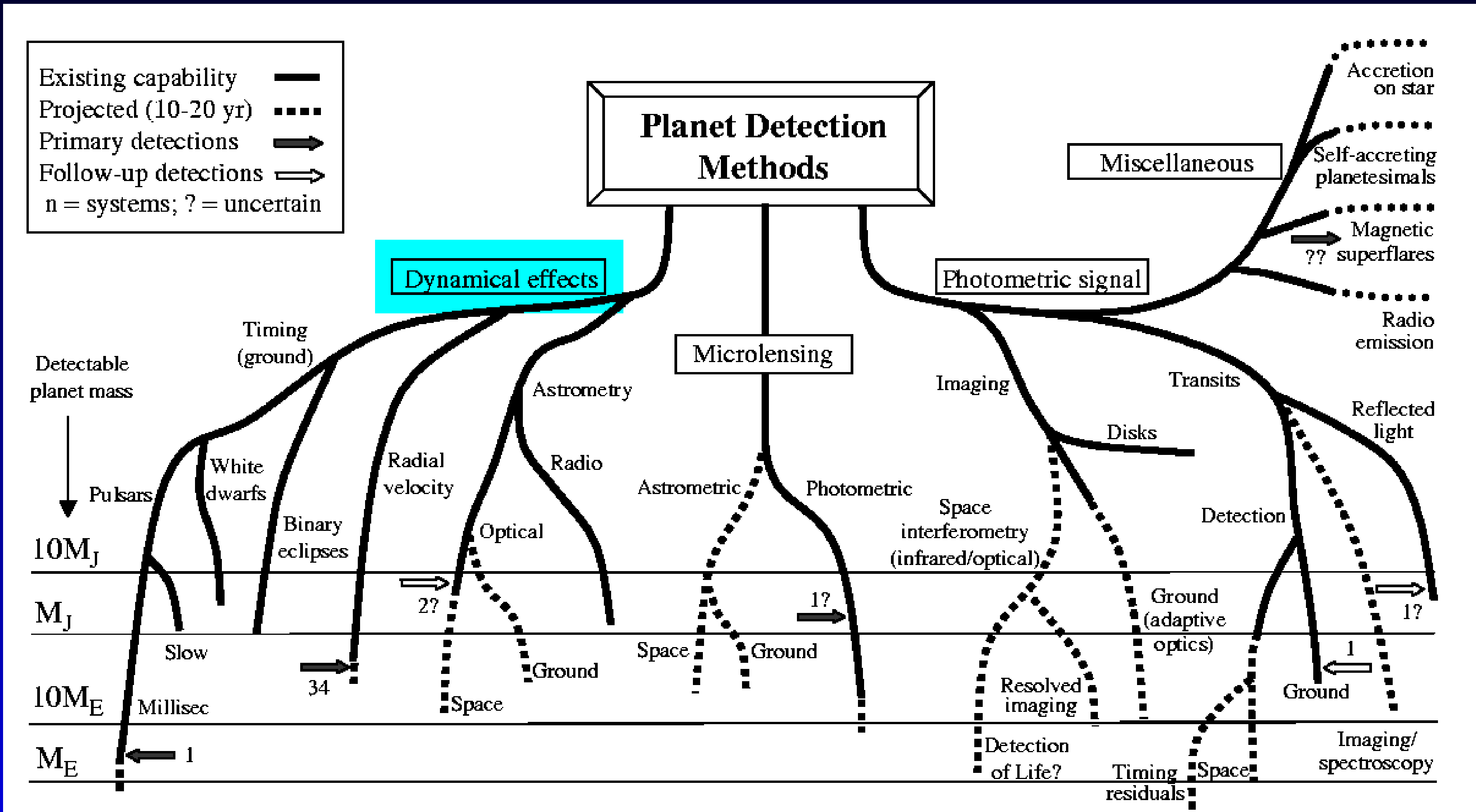
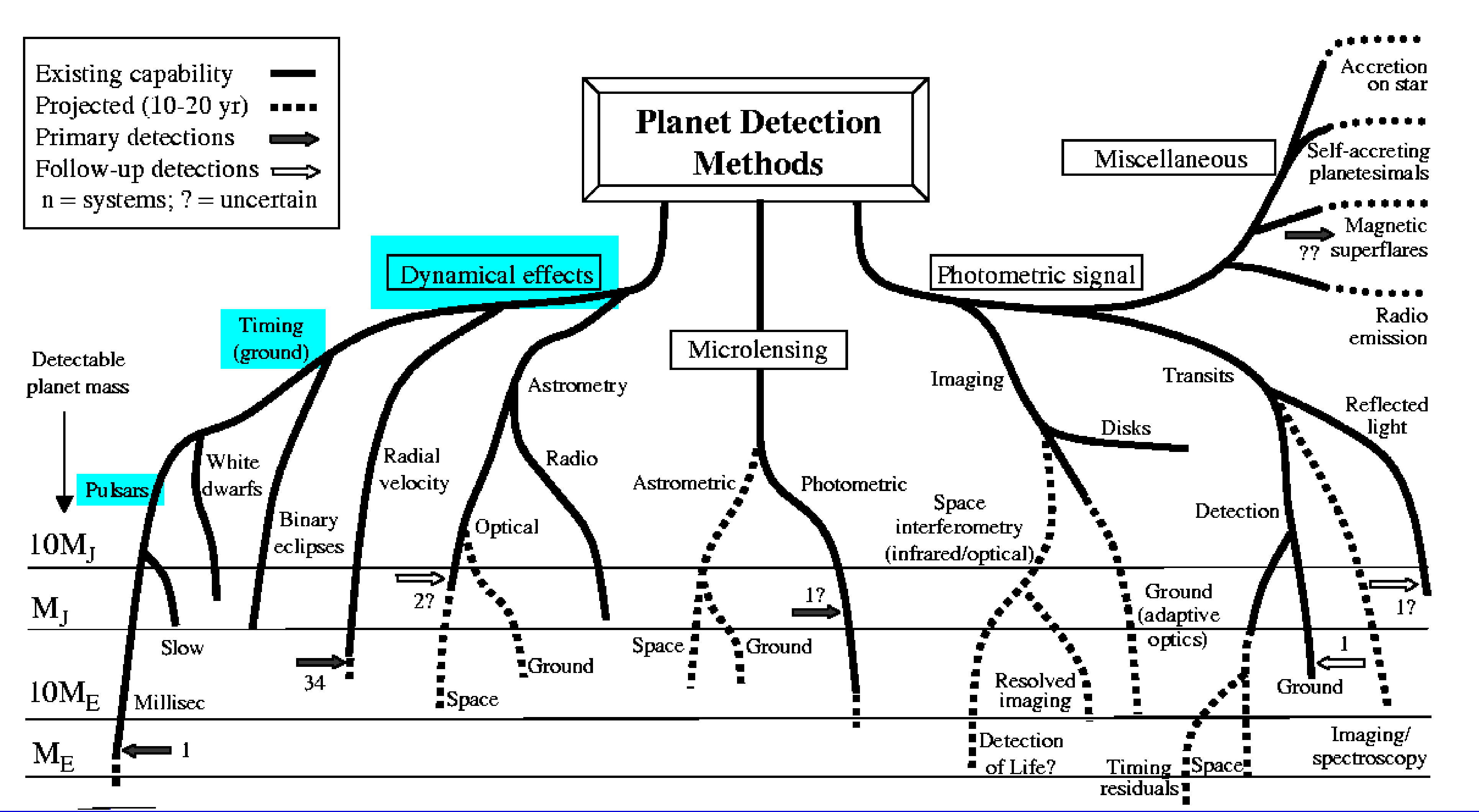


Image copyright M. Perryman

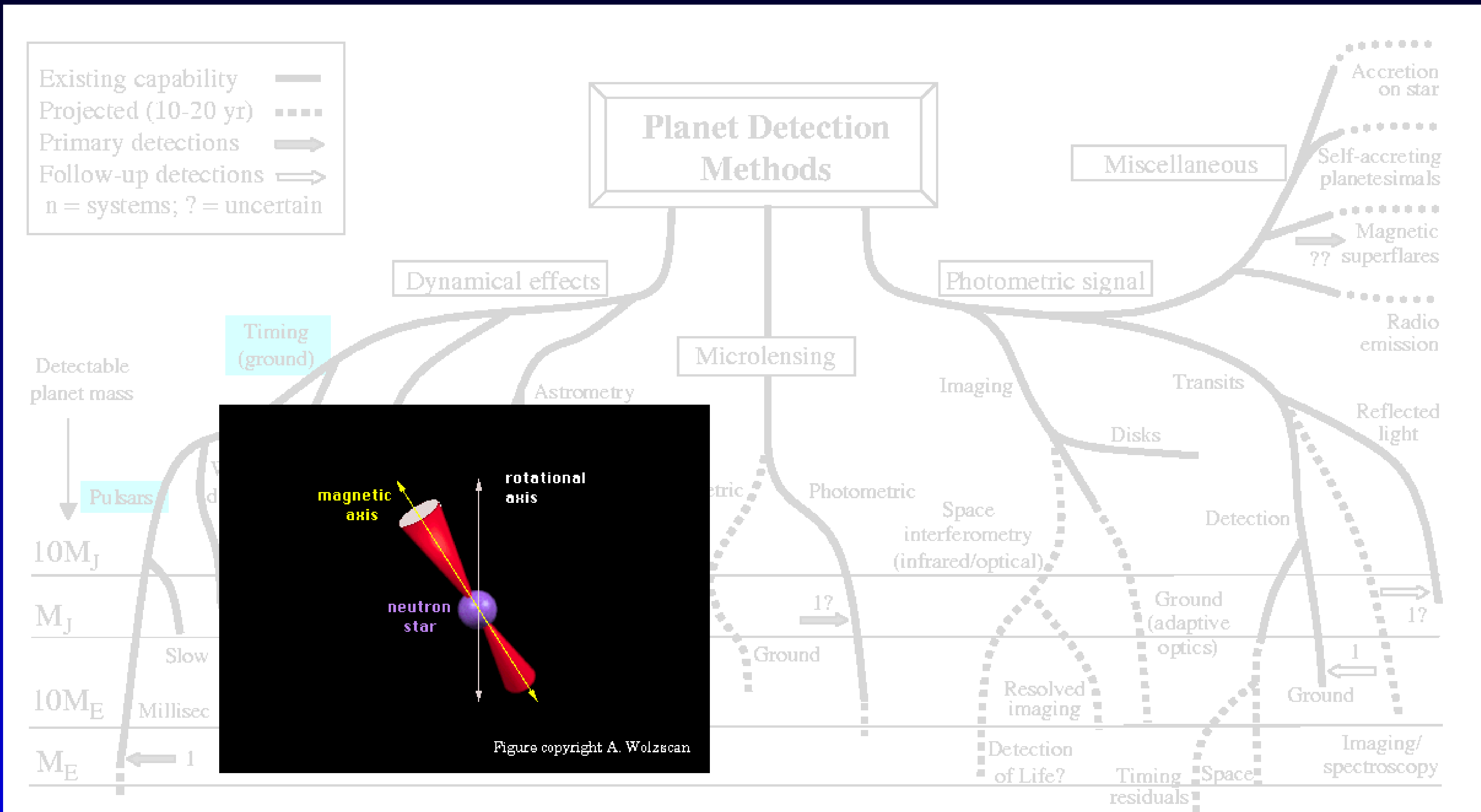
Detection techniques



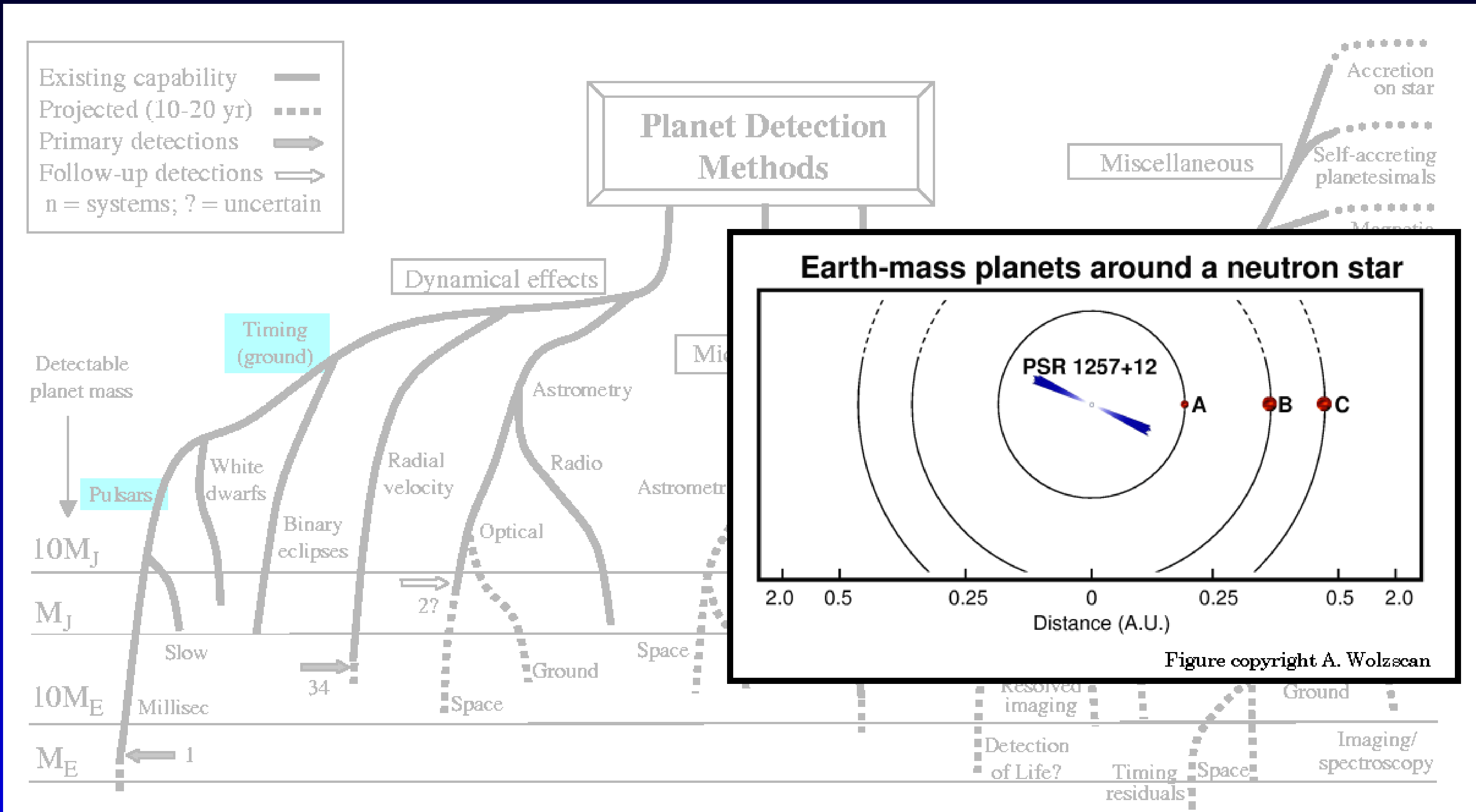
Detection techniques



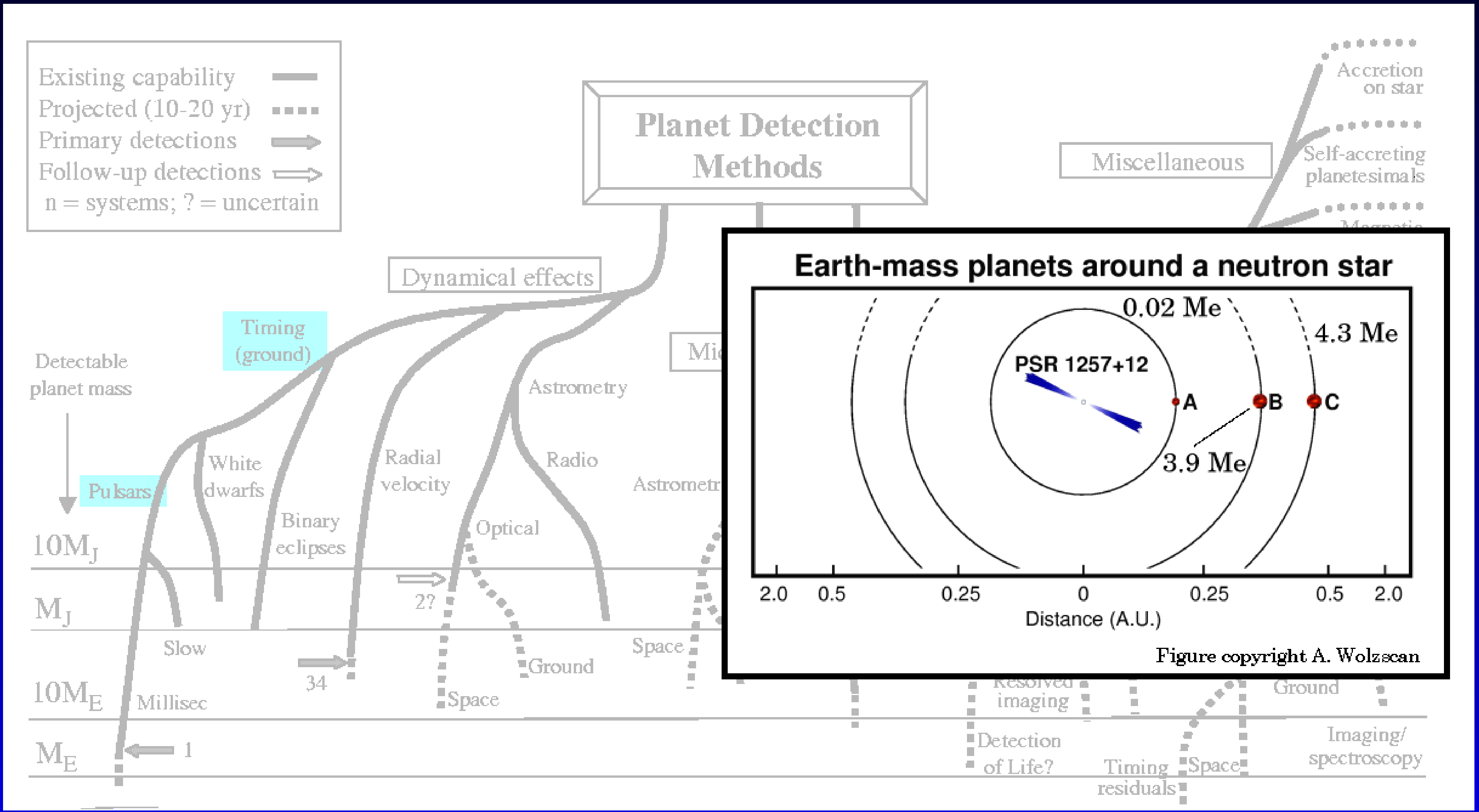
Detection techniques



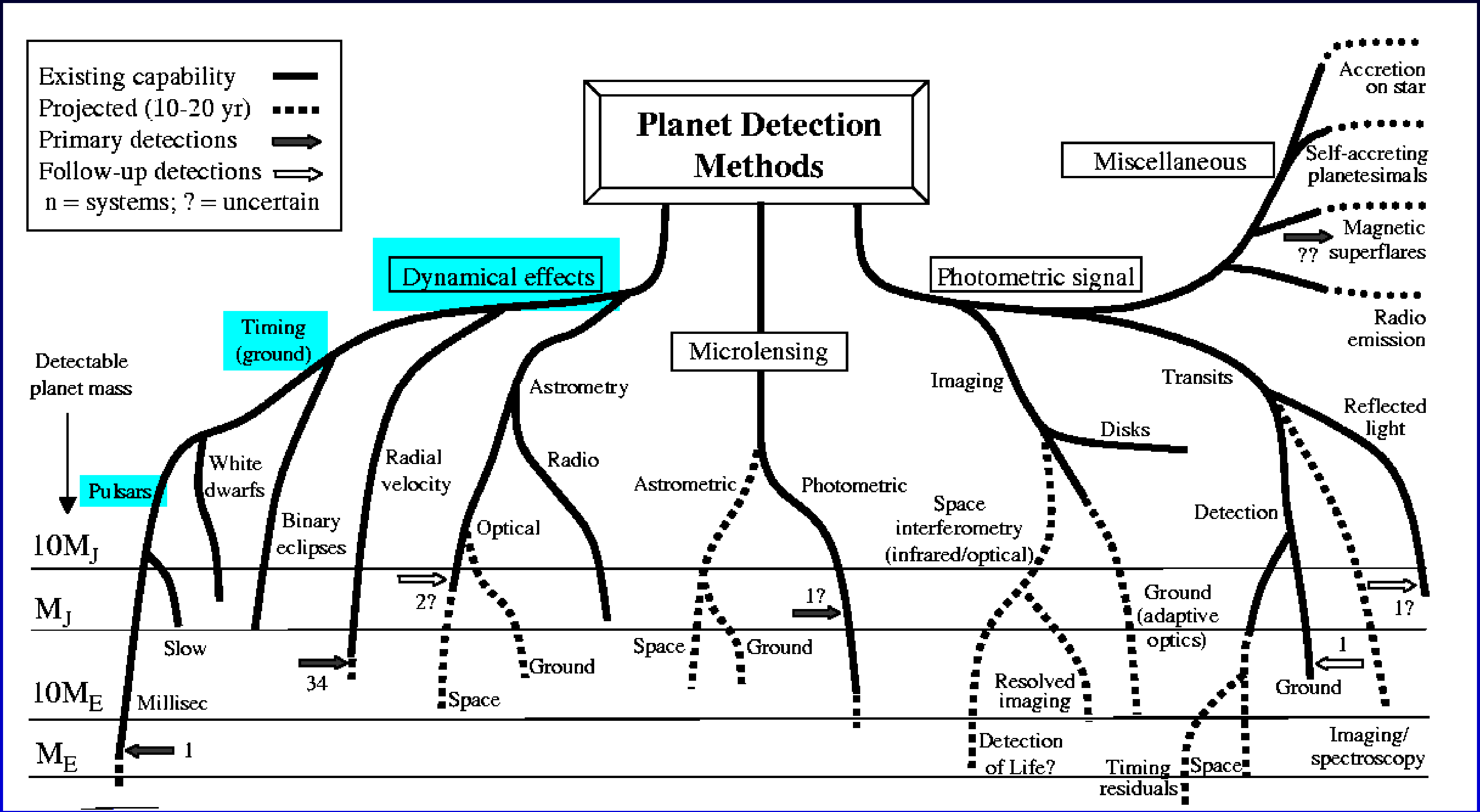
Detection techniques



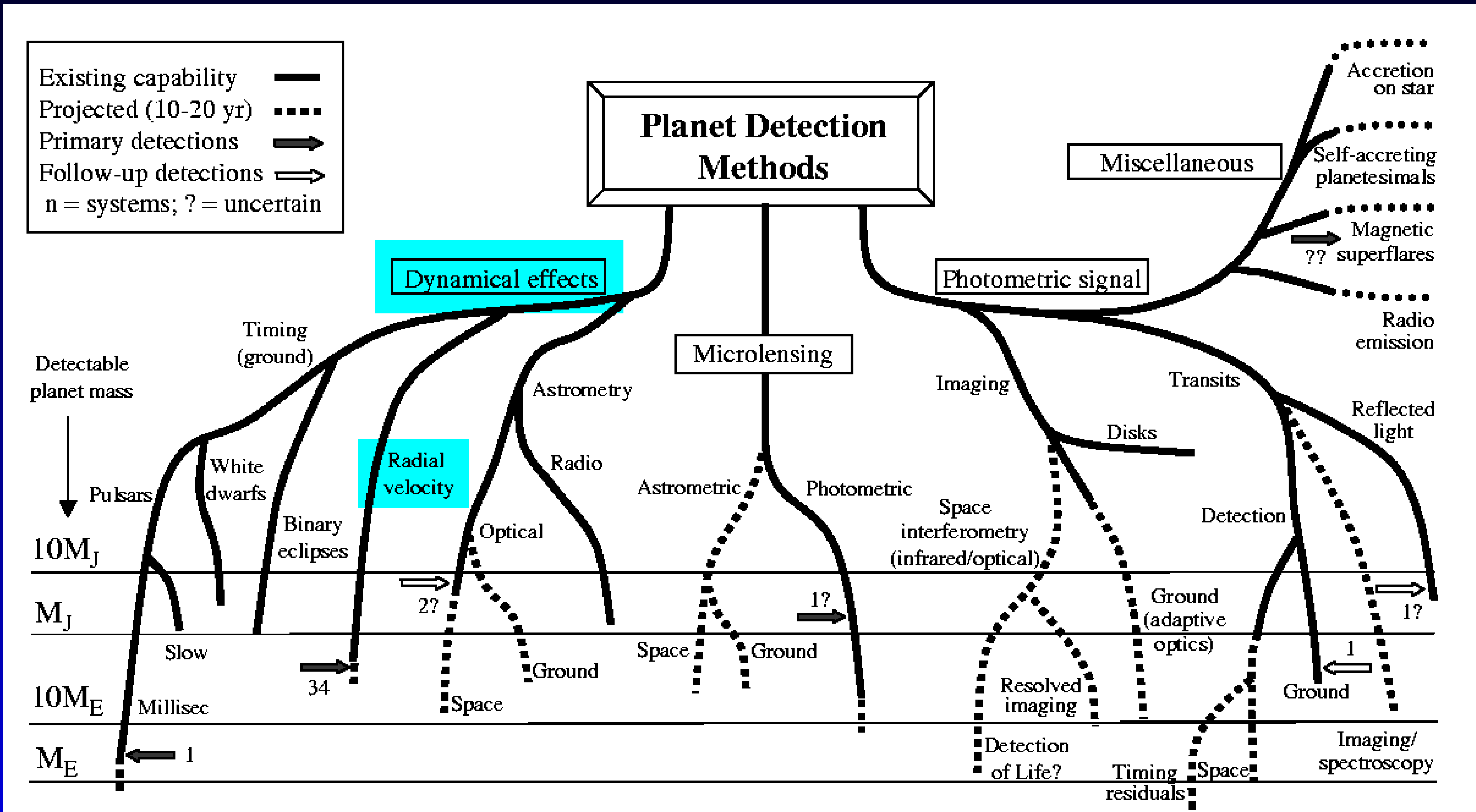
Detection techniques



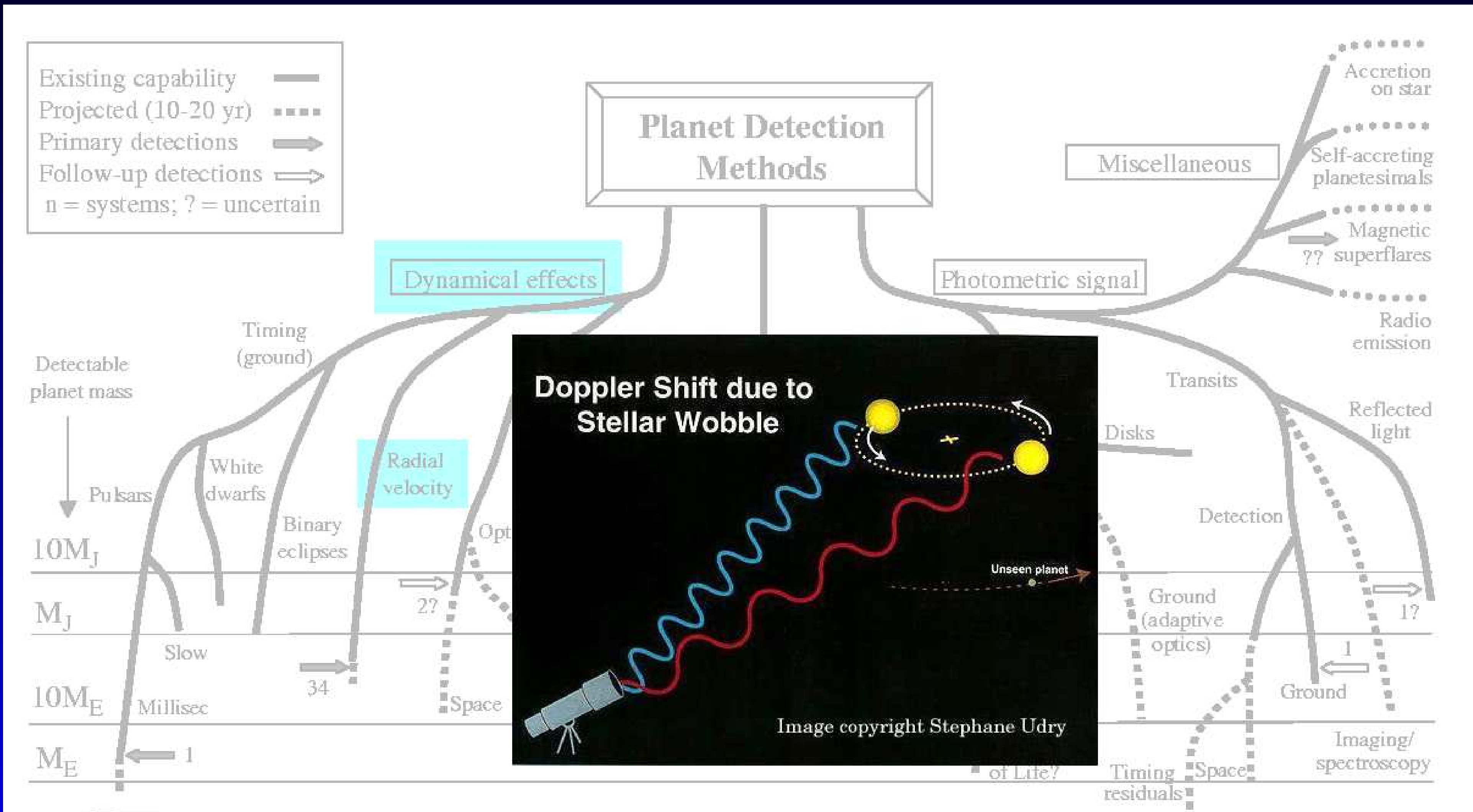
Detection techniques



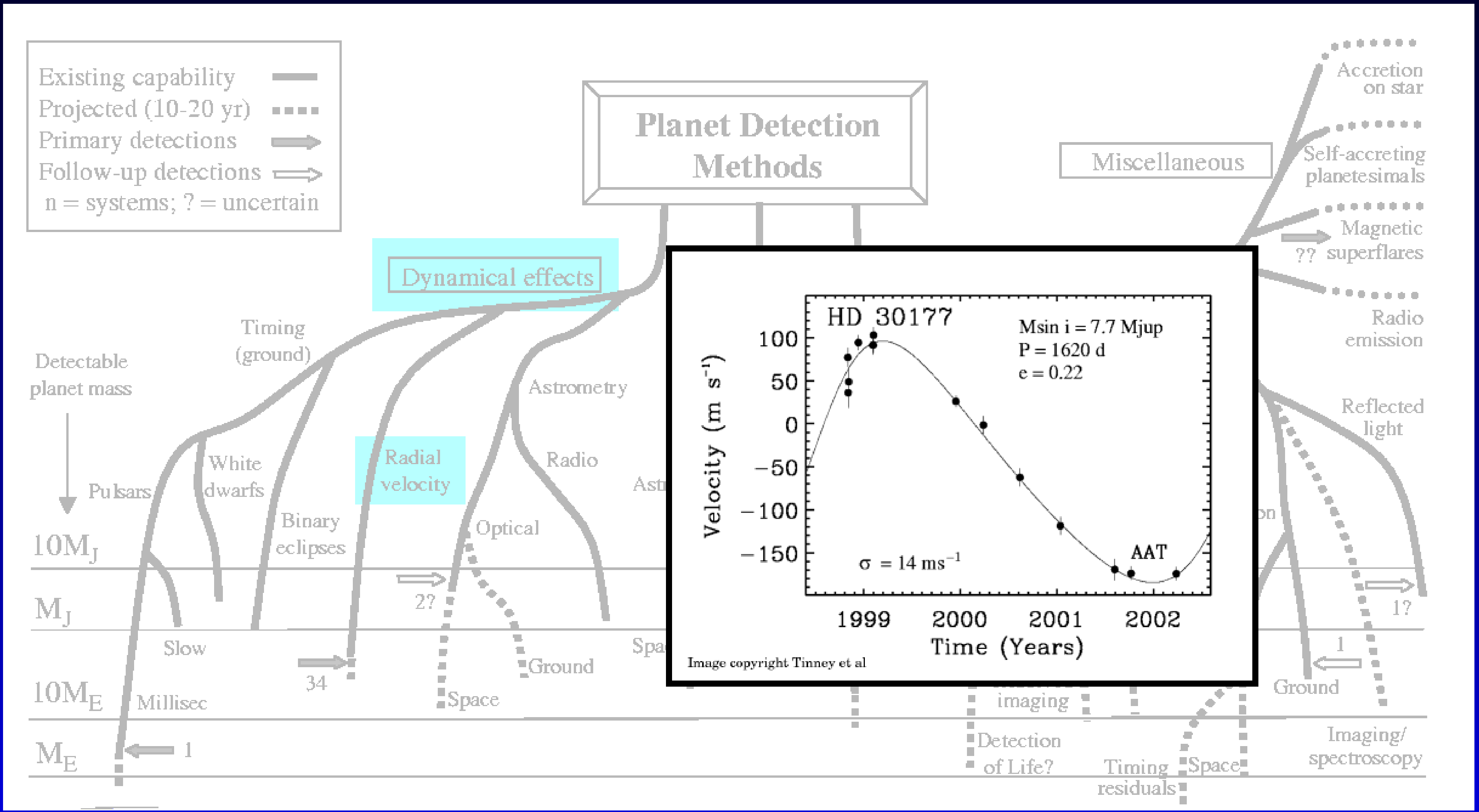
Detection techniques



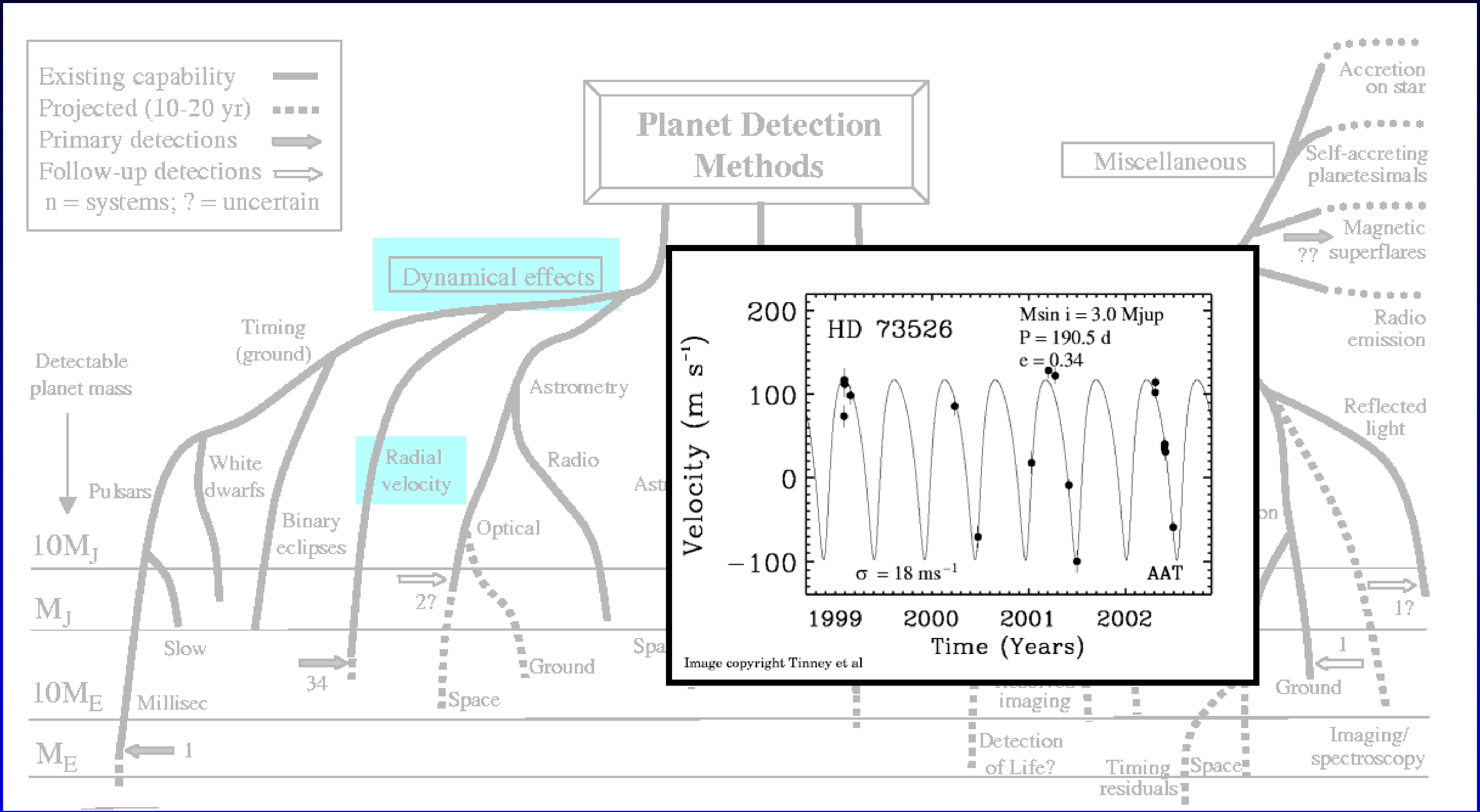
Detection techniques



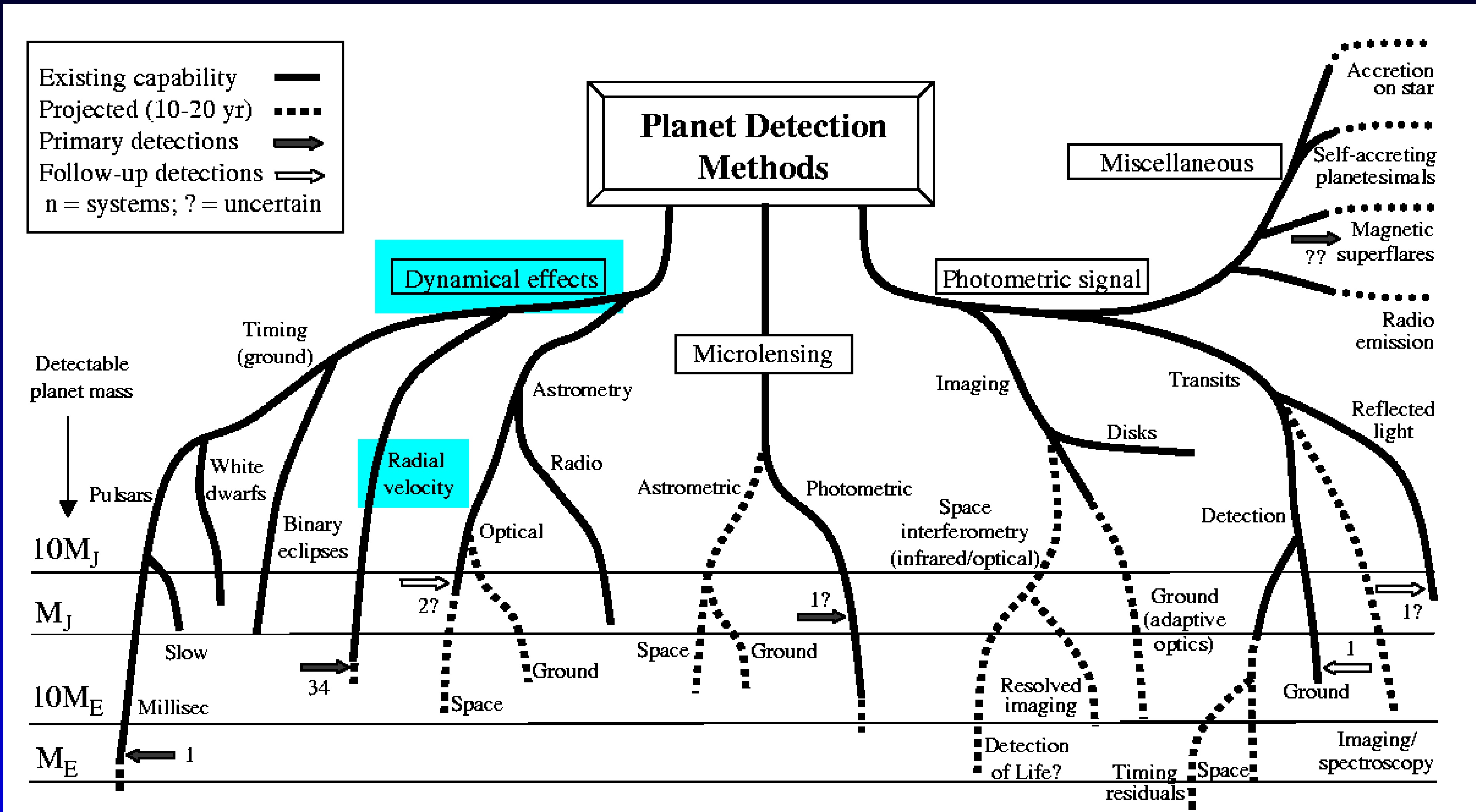
Detection techniques



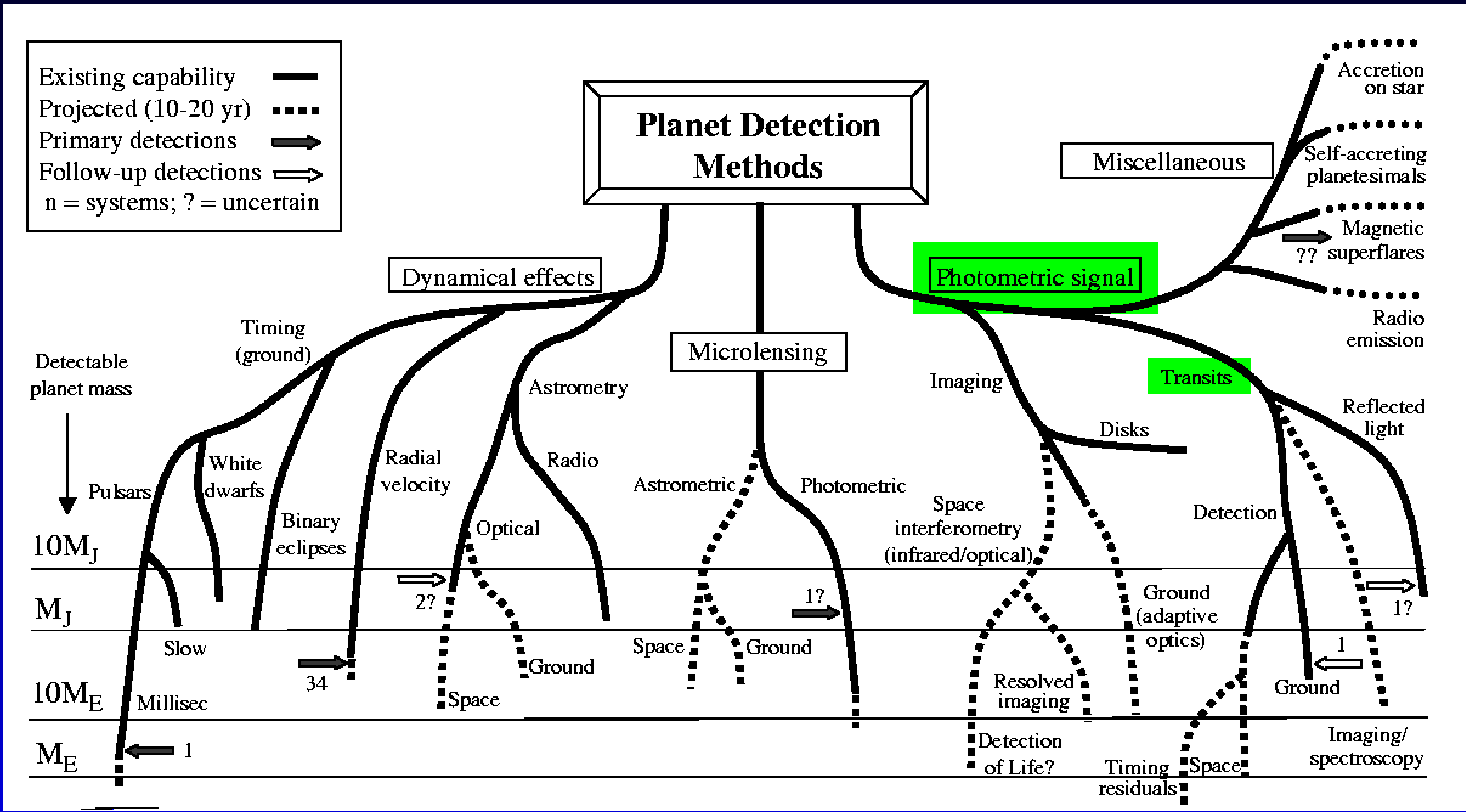
Detection techniques



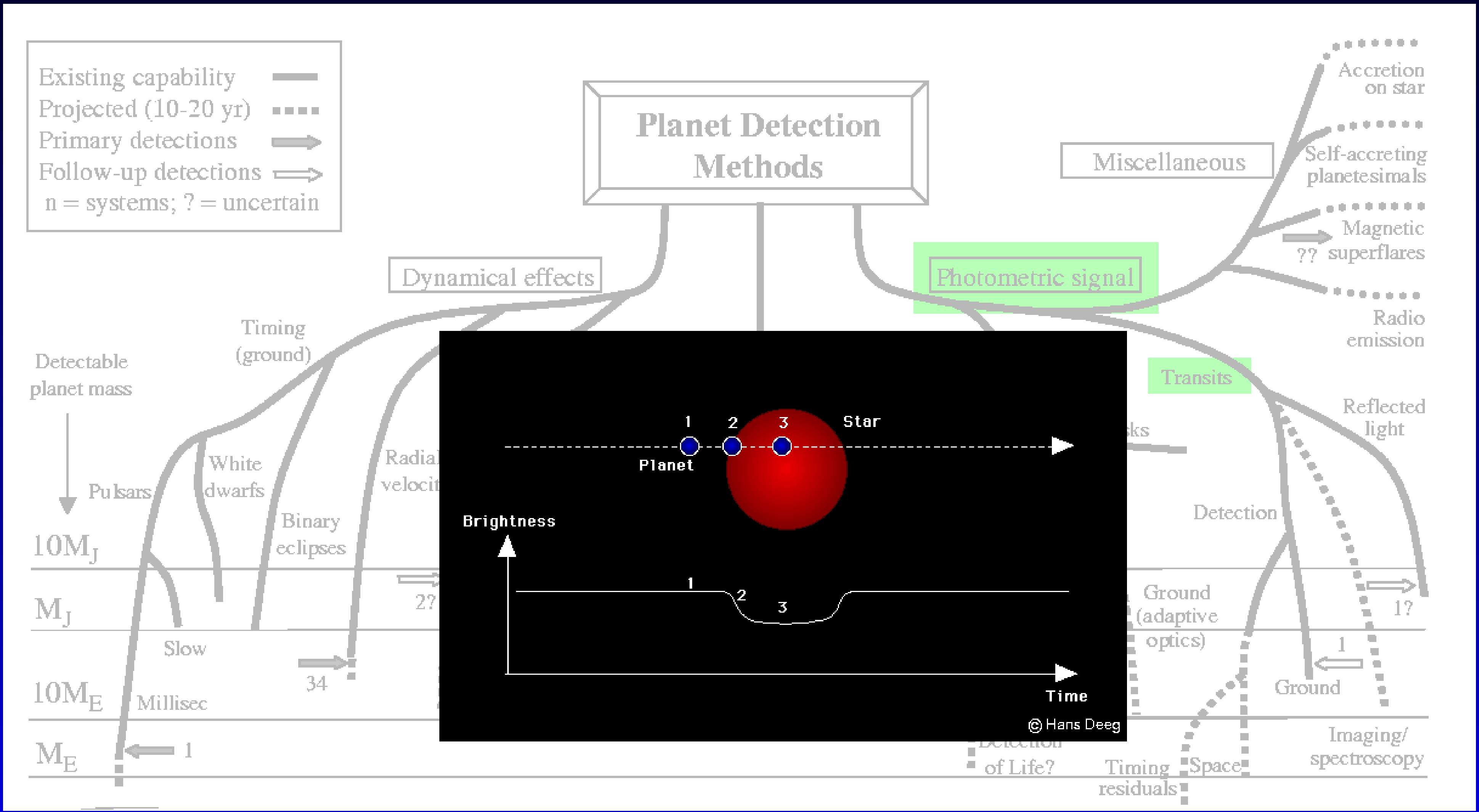
Detection techniques



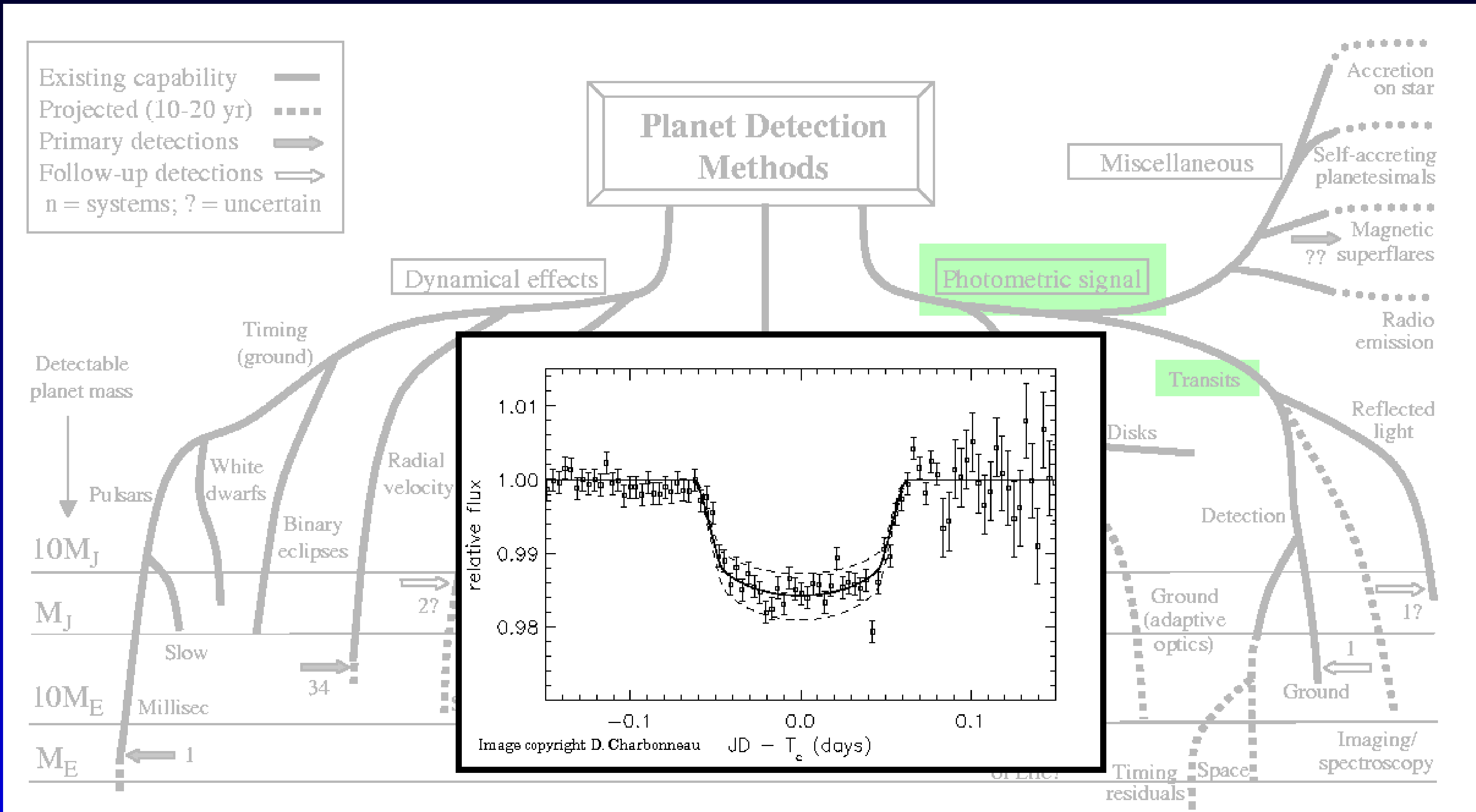
Detection techniques



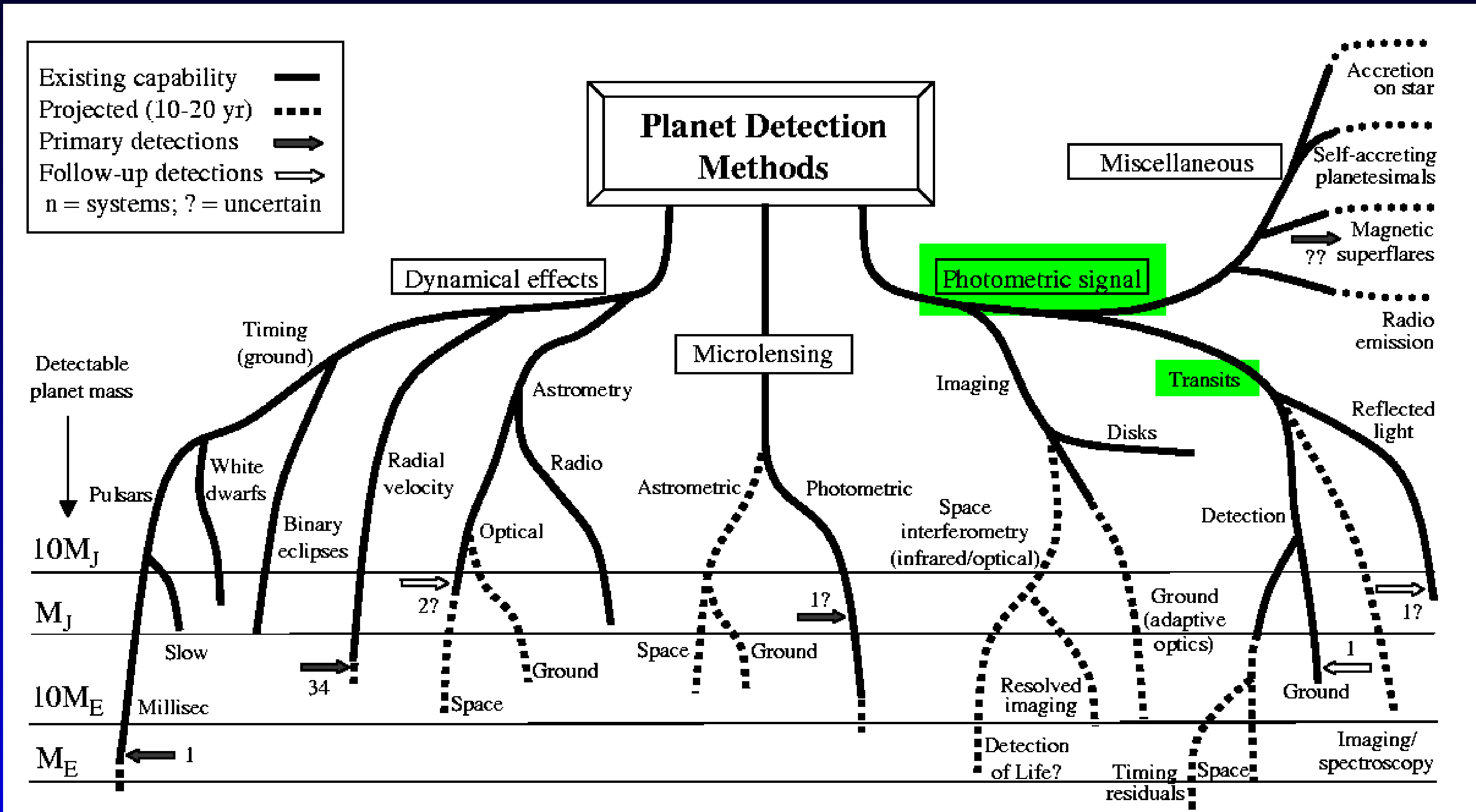
Detection techniques



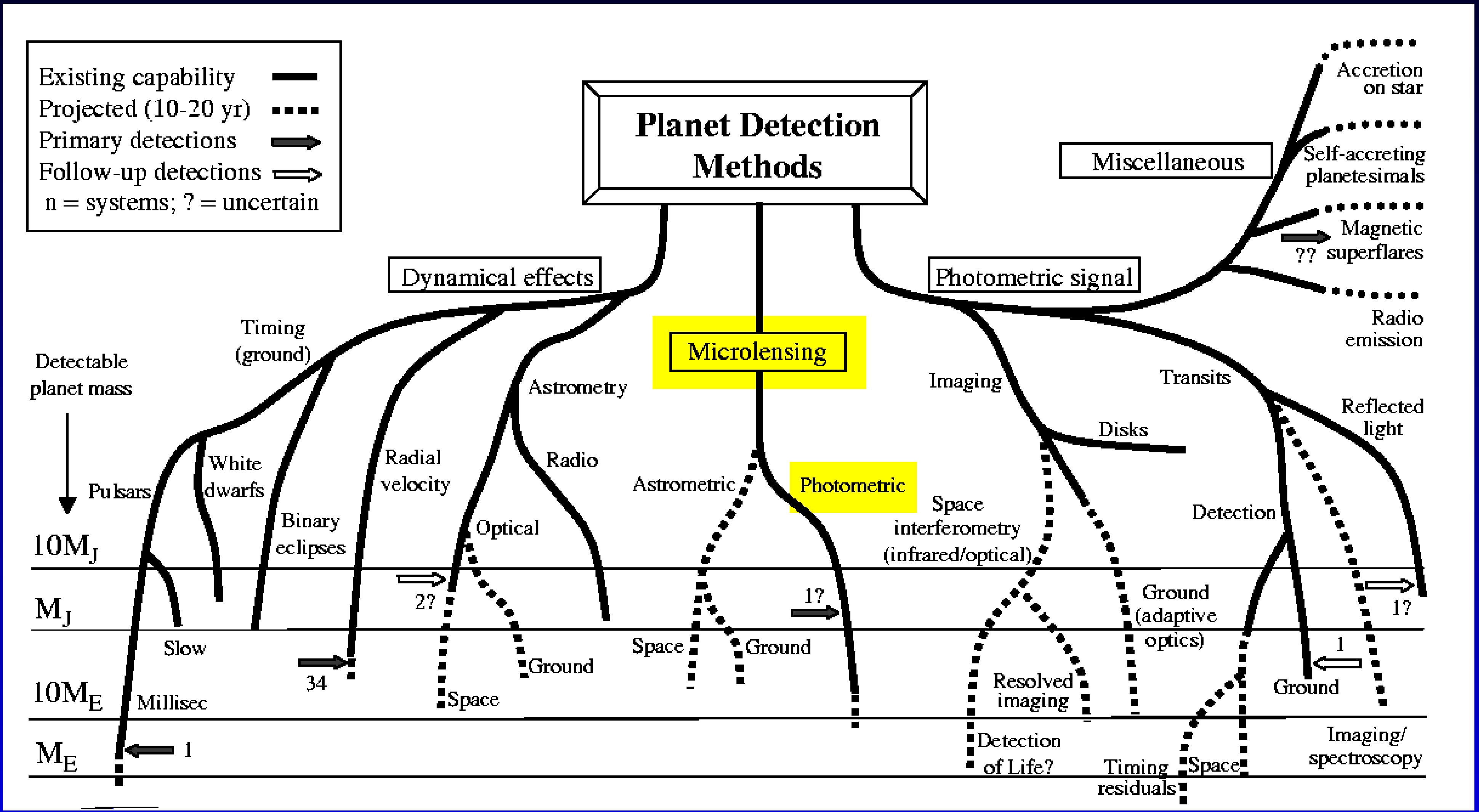
Detection techniques



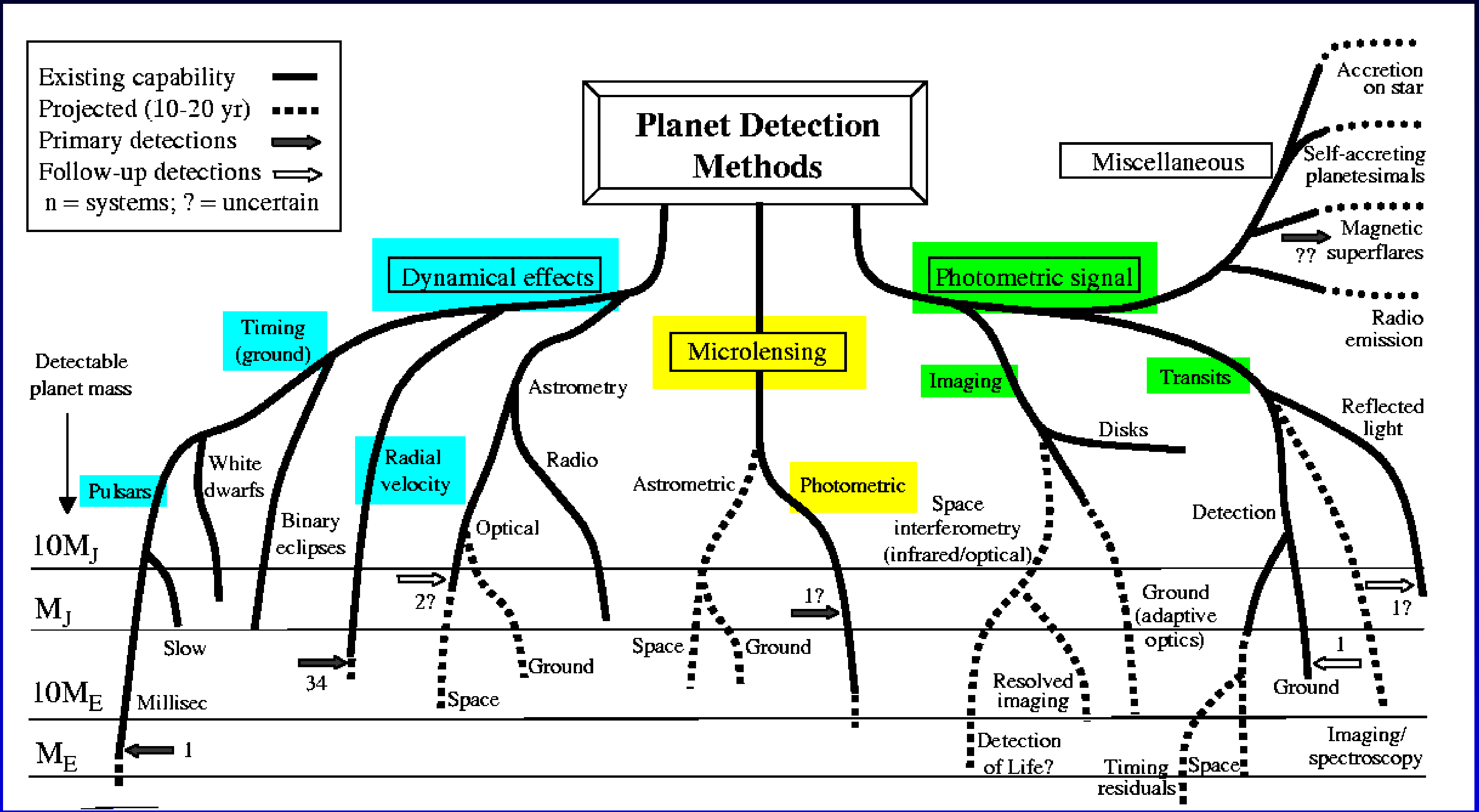
Detection techniques



Detection techniques



Detection techniques



Capability

Recall from the last lecture that microlensing is most sensitive to planets with projected orbit radii around $\sim R_E$.

Capability

Recall from the last lecture that microlensing is most sensitive to planets with projected orbit radii around $\sim R_E$.
This is an interesting regieme to probe, as the radial velocity and transit techniques are more effective at detecting planets close to their host star. However, they are extending their detection space as time baselines increase.

Capability

Recall from the last lecture that microlensing is most sensitive to planets with projected orbit radii around $\sim R_E$.
This is an interesting regieme to probe, as the radial velocity and transit techniques are more effective at detecting planets close to their host star. However, they are extending their detection space as time baselines increase.

Microlensing is currently detecting planets in a previously unreachable region of the planetary mass-radius space. Microlensing is returning detections of planets with masses approaching that of Earth.

Detection space

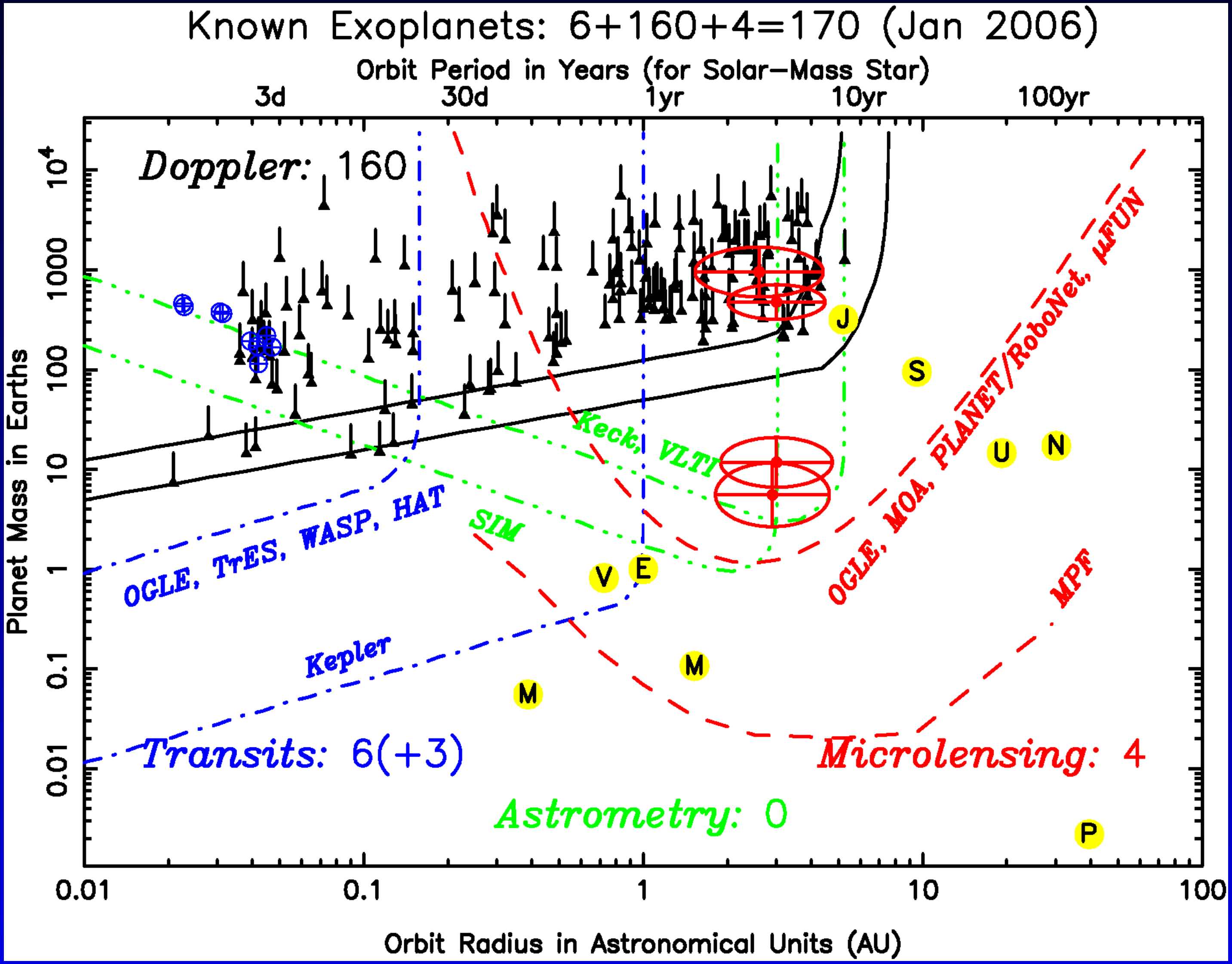


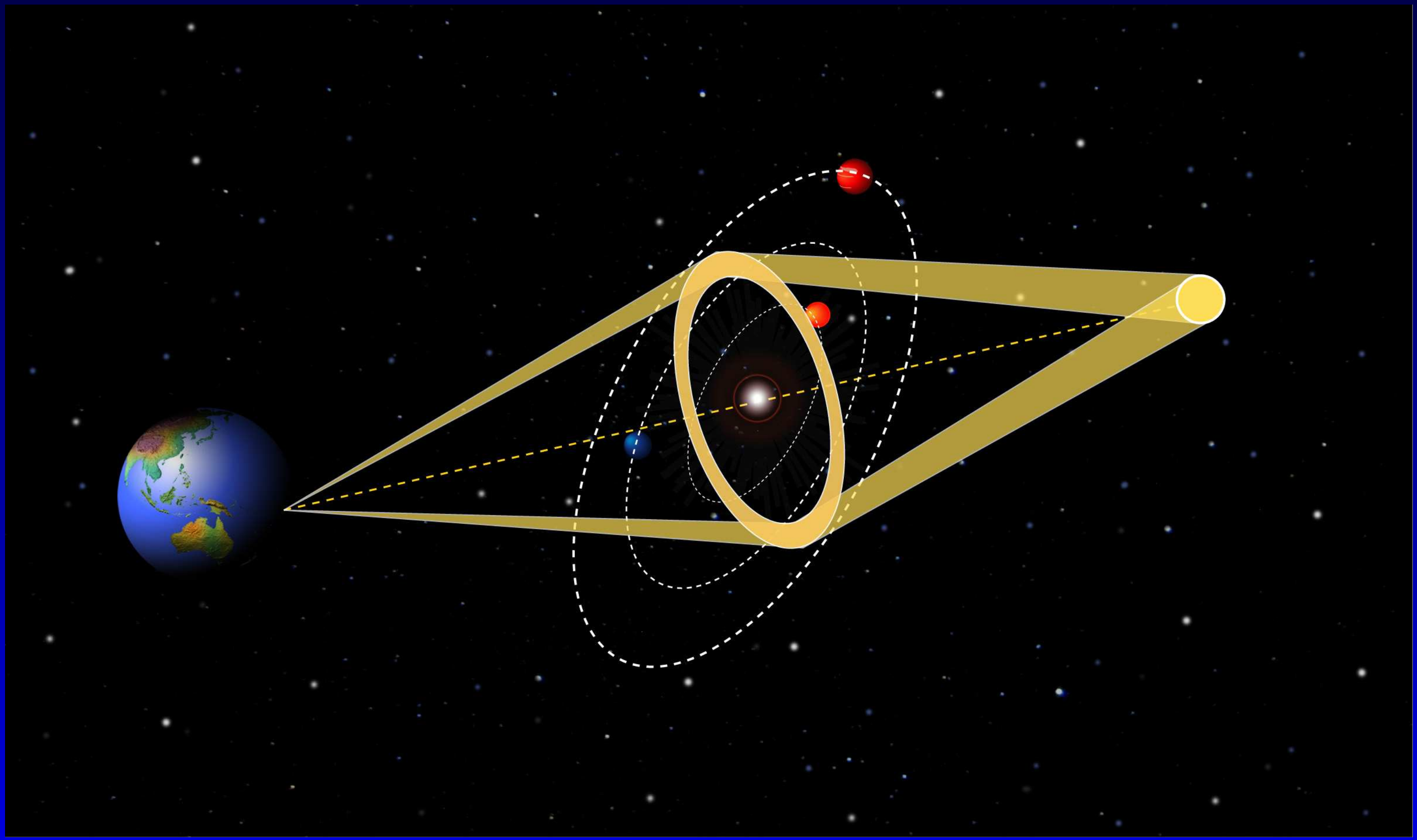
Figure courtesy K. Horne

Capability

The lensing effect is a “snapshot” of the lens system: all planets are detected at once.

Capability

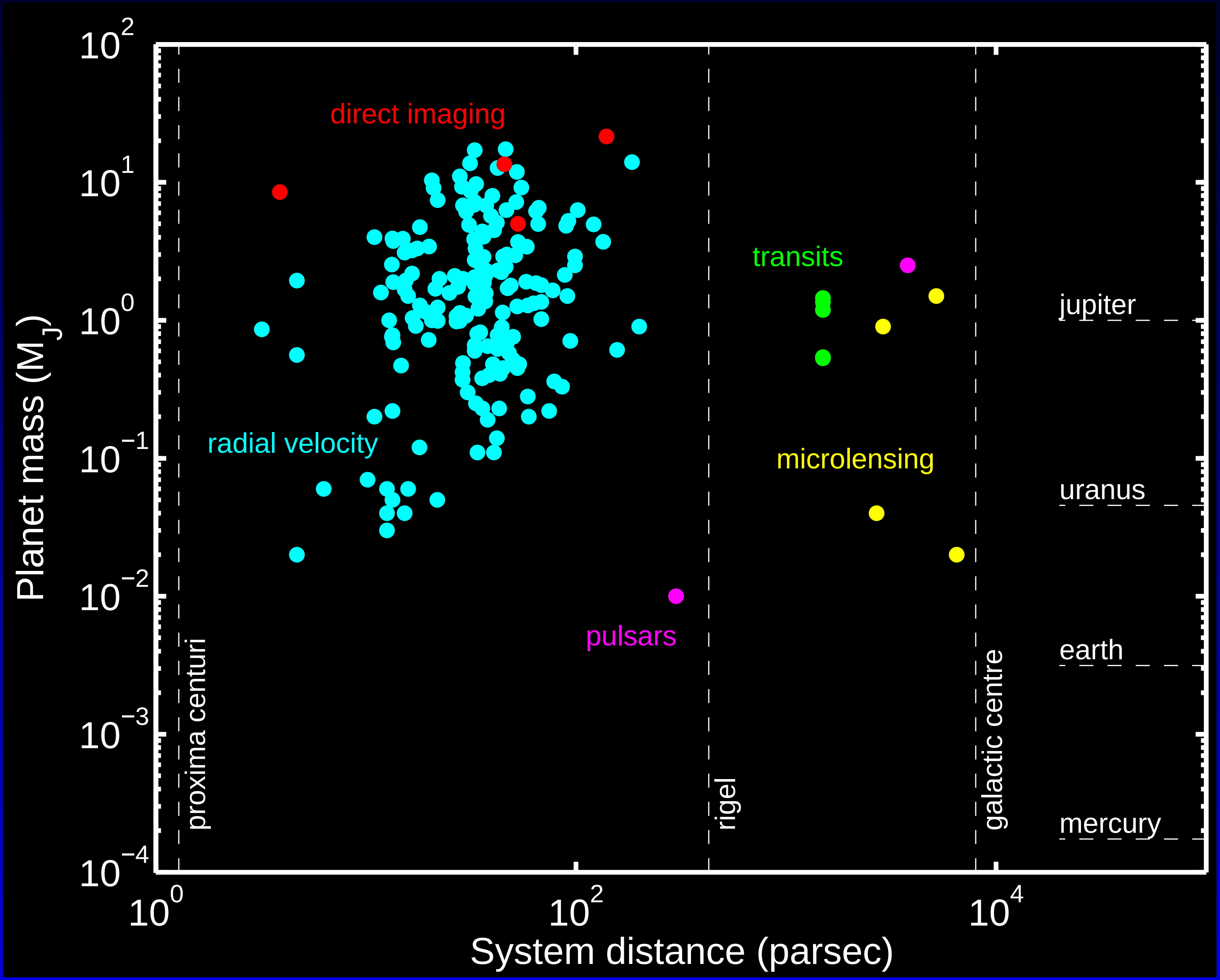
The lensing effect is a “snapshot” of the lens system: all planets are detected at once.



Capability

Distant planetary systems are being discovered.

Capability



direct imaging
radial velocity
transits
microlensing
pulsars
jupiter
uranus
earth
mercury
proxima centuri
rigel
galactic centre
Planet mass (M_J)
System distance (parsec)
10^2
10^1
10^0
10^{-1}
10^{-2}
10^{-3}
10^{-4}
10^0
10^2
10^4

Capability

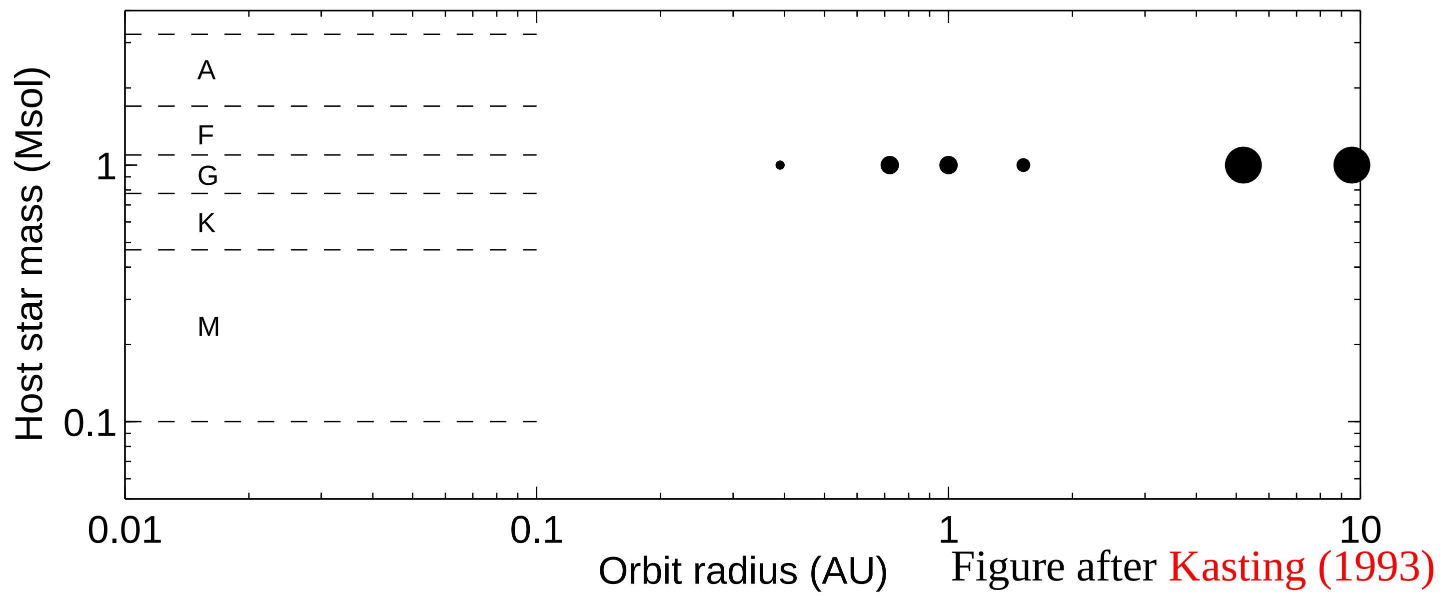
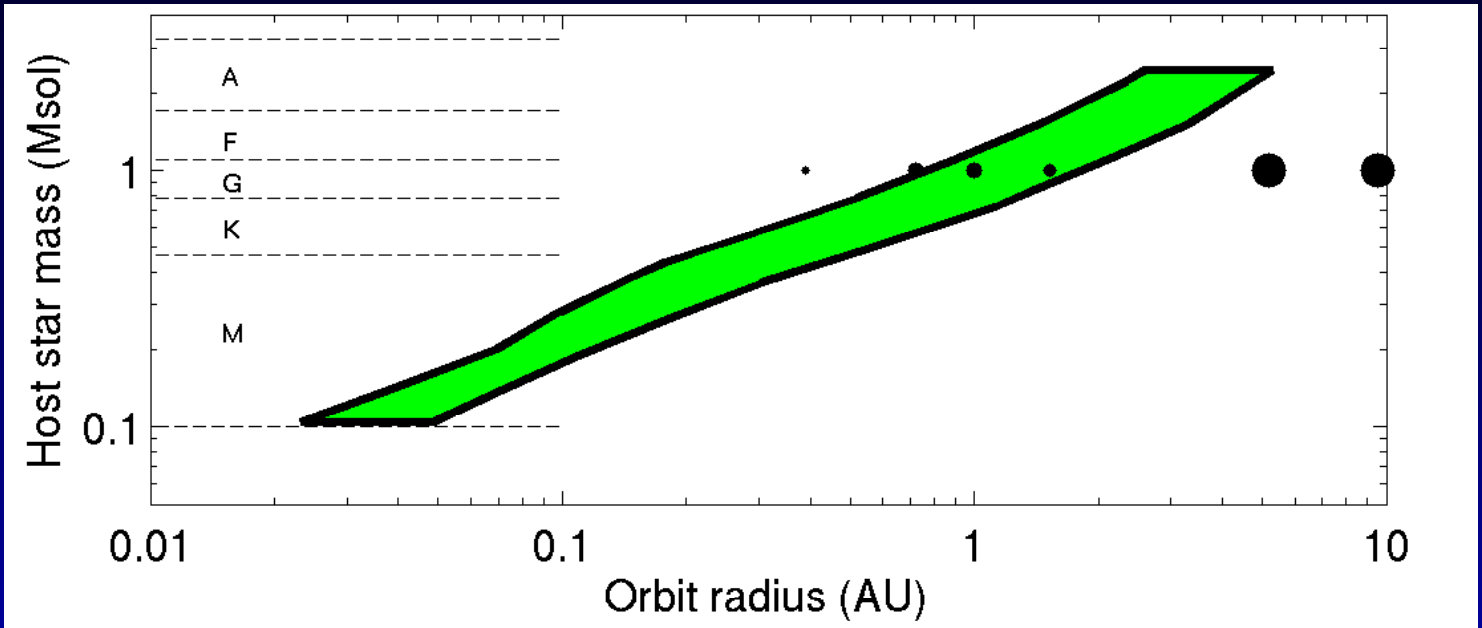
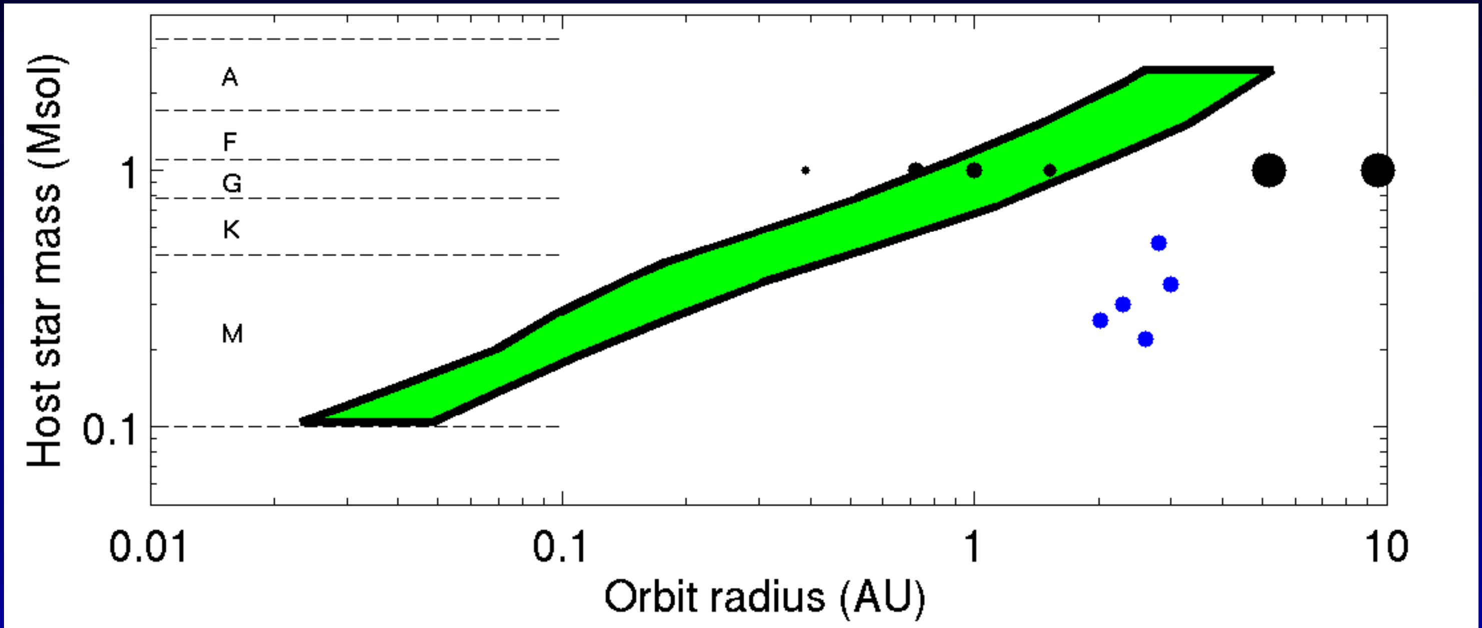


Figure after Kasting (1993)

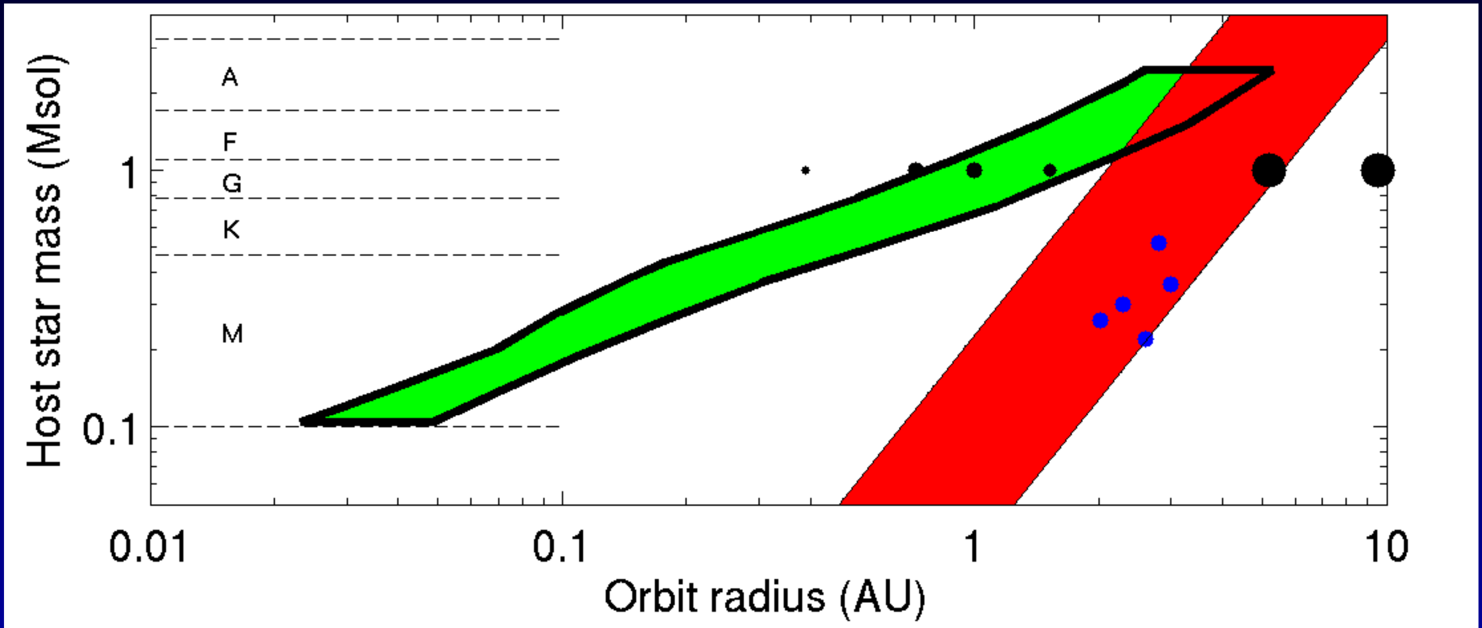
Capability



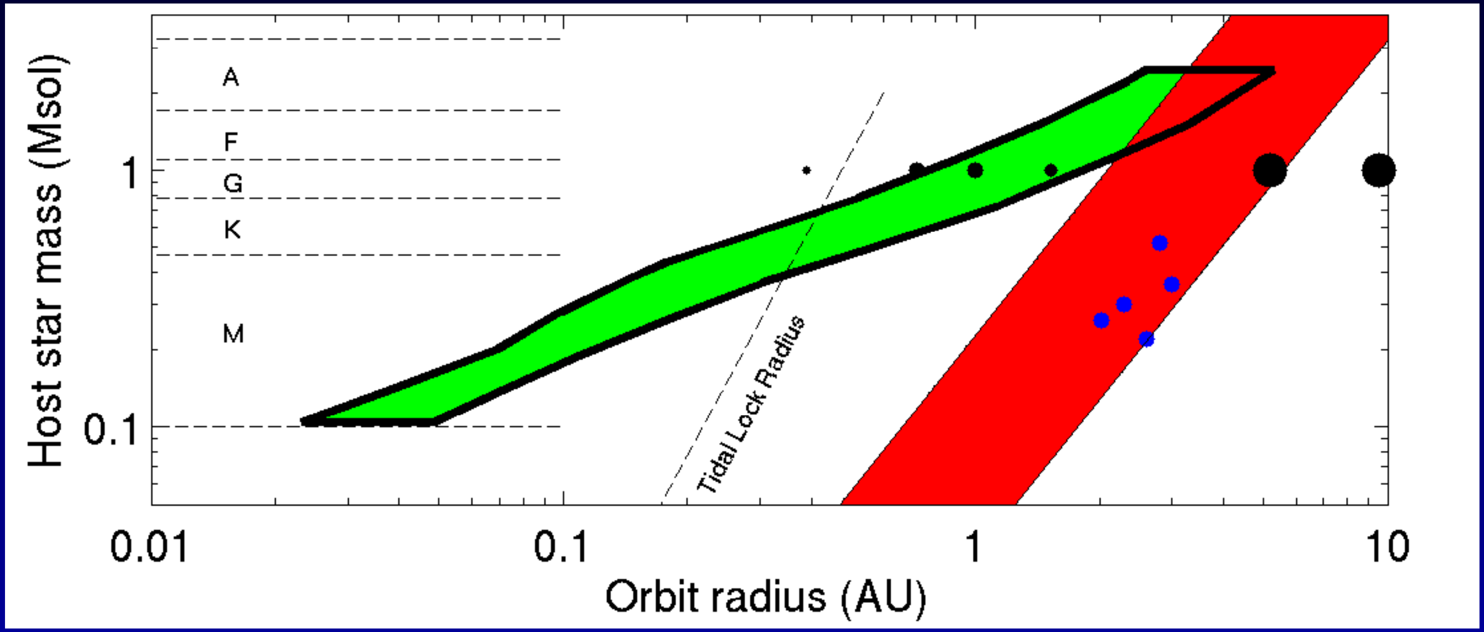
Capability



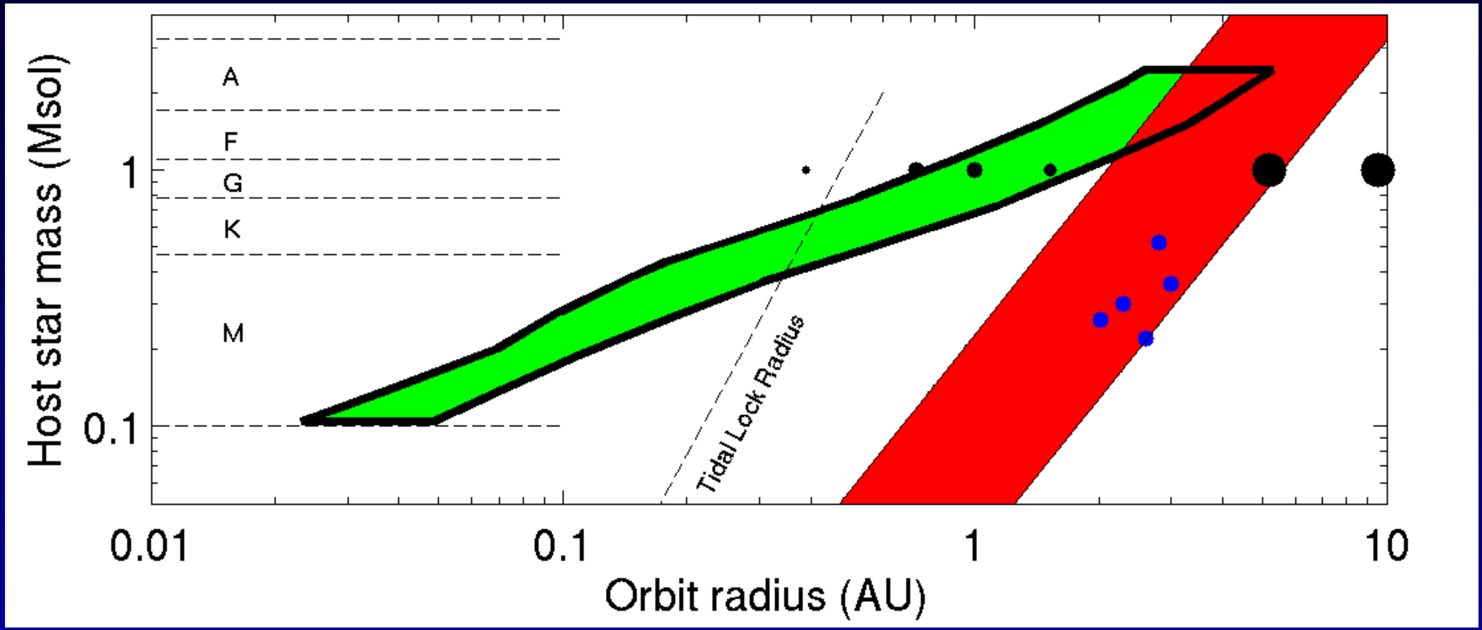
Capability



Capability

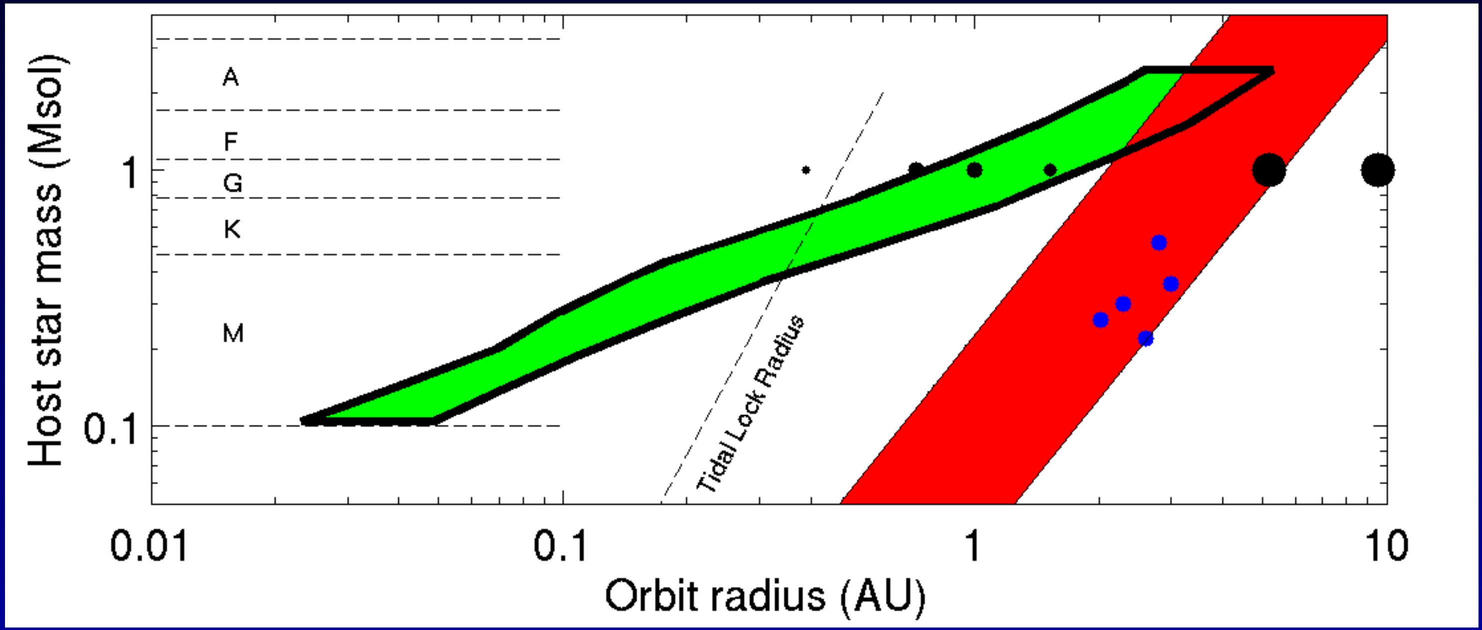

A
F
G
K
M
Tidal Lock Radius
Host star mass (Msol)
1
0.1
0.01
0.1
1
10
Orbit radius (AU)

Capability



Ground-based microlensing is relatively insensitive to habitable planets, as the most-likely lens stars are cool M-dwarfs.

Capability



Ground-based microlensing is relatively insensitive to habitable planets, as the most-likely lens stars are cool M-dwarfs. However, a space telescope such as the proposed Microlensing Planet Finder (Bennett, 2006) will be much more likely to detect habitable planets via microlensing, including moons, if present.

Detection

Given a microlensing event, our default assumption is that it is due to a single lens object.

Detection

Given a microlensing event, our default assumption is that it is due to a single lens object.

However:

In studies of nearby solar type stars, Abt & Levy (1976) found 58% of the G dwarf primary stars had one or more stellar companions(see also Abt, 1987). Duquennoy & Marcy (1991) find a similar fraction of 57%. The multiplicity amongst M dwarfs is slightly less, at 42% (Fischer & Marcy, 1992).

Detection

Given a microlensing event, our default assumption is that it is due to a single lens object.

Any anomaly seen in a microlensing event attributable to a non-singular lens mass is based on the significance of the departure from a single lens fit to the lightcurve.

Detection

Given a microlensing event, our default assumption is that it is due to a single lens object.

Any anomaly seen in a microlensing event attributable to a non-singular lens mass is based on the significance of the departure from a single lens fit to the lightcurve.

The significance of any deviation is going to be strongly dependent on the quality of the photometry and the frequency and consistency of observations.

Detection

Given a microlensing event, our default assumption is that it is due to a single lens object.

Any anomaly seen in a microlensing event attributable to a non-singular lens mass is based on the significance of the departure from a single lens fit to the lightcurve.

The significance of any deviation is going to be strongly dependent on the quality of the photometry and the frequency and consistency of observations.

Most common “goodness-of-fit” paramater is

$$\chi^2 = \sum_{i=1}^{N} \frac{(y_i(t) - \hat{y}_i(t))^2}{\sigma_i^2}$$

Detection

We can use the difference in χ^2 between a single lens fit and a binary lens fit as our detection criteria.

Detection

We can use the difference in χ^2 between a single lens fit and a binary lens fit as our detection criteria.

$$\Delta\chi^2 = \chi^2_{\mathrm{S}} - \chi^2_{\mathrm{P}}$$

where χ^2_{S} and χ^2_{P} are the χ^2 values for a fit to a single lens and binary(planetary) lens fit respectively.

Detection

We can use the difference in χ^2 between a single lens fit and a binary lens fit as our detection criteria.

$$\Delta\chi^2 = \chi^2_{\mathrm{S}} - \chi^2_{\mathrm{P}}$$

where χ^2_{S} and χ^2_{P} are the χ^2 values for a fit to a single lens and binary(planetary) lens fit respectively.

Typical detection threshold values are $\Delta\chi^2 = 60$ – 100, corresponding to significance levels of $\sim 6.3\sigma$ – 10σ.

Detection

We can use the difference in χ^2 between a single lens fit and a binary lens fit as our detection criteria.

$$\Delta\chi^2 = \chi^2_{\rm S} - \chi^2_{\rm P}$$

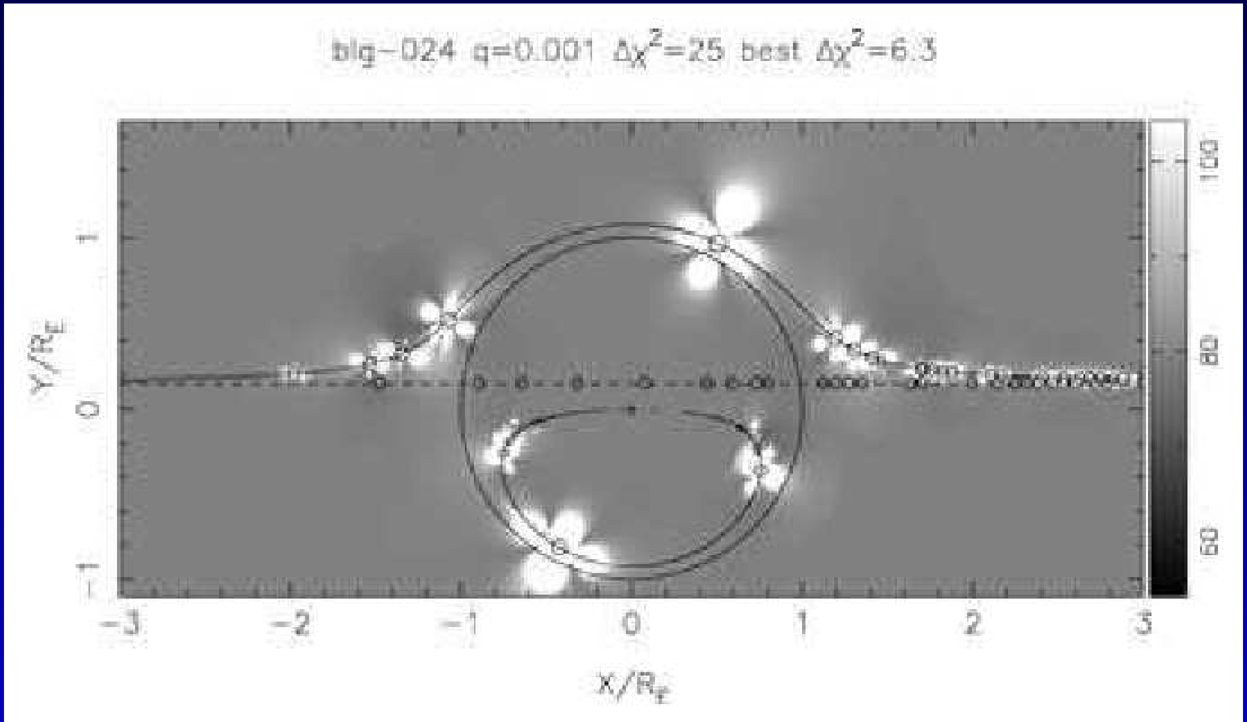
where $\chi^2_{\rm S}$ and $\chi^2_{\rm P}$ are the χ^2 values for a fit to a single lens and binary(planetary) lens fit respectively. Typical detection threshold values are $\Delta\chi^2 = 60$ – 100, corresponding to significance levels of $\sim 6.3\sigma$ – 10σ. $\Delta\chi^2$ as a detection criterion is sensitive to, amongst other things, noisy data and photometric errors.

Detection

Can produce detection zone maps per observed event, e.g. Snodgrass, Horne & Tsapras (2004):

Detection

Can produce detection zone maps per observed event, e.g. Snodgrass, Horne & Tsapras (2004):

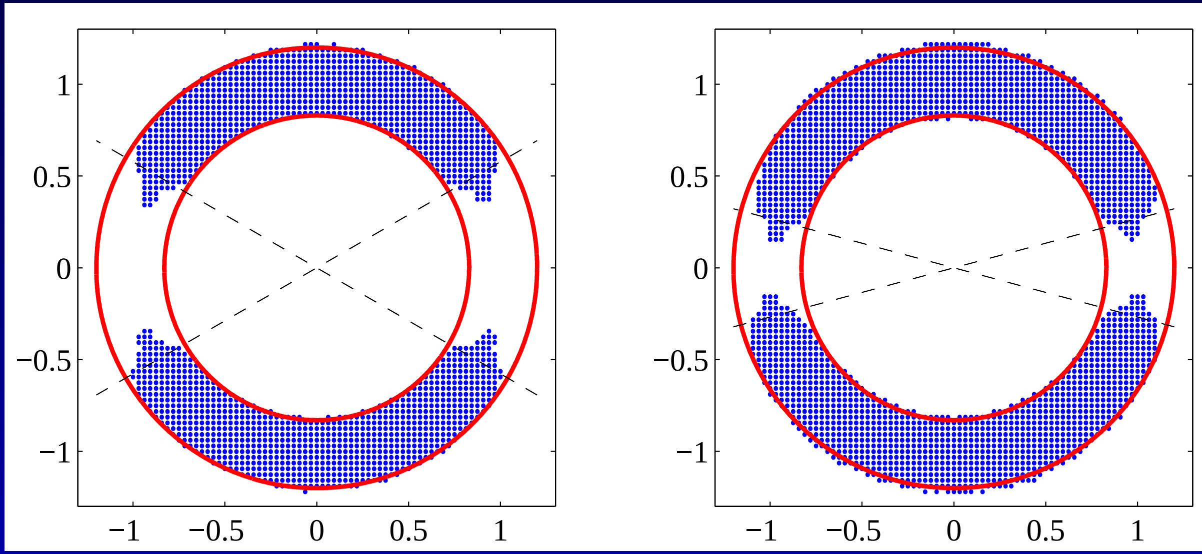


Detection

Or for given theoretical events, e.g.
(Rattenbury, 2002)

Detection

Or for given theoretical events, e.g. (Rattenbury, 2002)



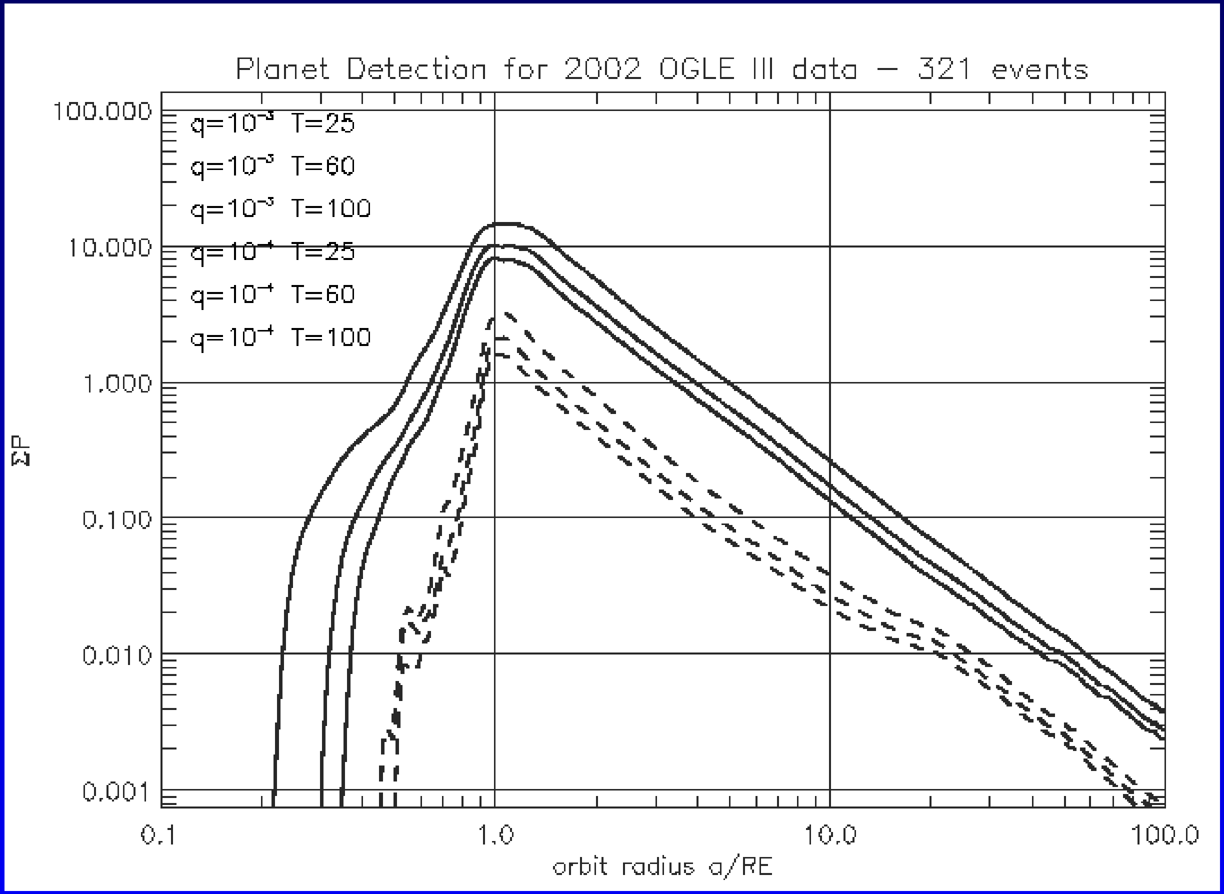
Detection limit maps for an Earth mass planet orbiting a $0.3M_\odot$ lens star. The axes are in units of R_E. The I band magnitude at maximum is 15, and the maximum amplification is 100. The χ^2 map on the left was generated using observation light curves comprised of 301 points over the interval $[-\frac{1}{2}t_{\mathrm{FWHM}}, \frac{1}{2}t_{\mathrm{FWHM}}]$. The right map was created using light curves comprised of 601 points over the interval $[-t_{\mathrm{FWHM}}, t_{\mathrm{FWHM}}]$.

Detection

Snodgrass, Horne & Tsapras (2004) also computed the detection zones for a range of planets and predicted the number of planets detected assuming the OGLE-III experiment characteristics.

Detection

Snodgrass, Horne & Tsapras (2004) also computed the detection zones for a range of planets and predicted the number of planets detected assuming the OGLE-III experiment characteristics.



Detection

$\Delta\chi^2$ is fine for theoretical investigations, or formal comparison between competing models.

Detection

$\Delta\chi^2$ is fine for theoretical investigations, or formal comparison between competing models.

But as a first detection criterion, can be easily fooled in practice with real observational data.

Detection

$\Delta\chi^2$ is fine for theoretical investigations, or formal comparison between competing models.

But as a first detection criterion, can be easily fooled in practice with real observational data.

Use is often made of some other detection criteria such as:

Detection

$\Delta\chi^2$ is fine for theoretical investigations, or formal comparison between competing models.

But as a first detection criterion, can be easily fooled in practice with real observational data.

Use is often made of some other detection criteria such as:

- Point-by-point significance - e.g 3 consecutive points deviating by $\geq 3\sigma$

Detection

$\Delta\chi^2$ is fine for theoretical investigations, or formal comparison between competing models.

But as a first detection criterion, can be easily fooled in practice with real observational data.

Use is often made of some other detection criteria such as:

- Point-by-point significance - e.g 3 consecutive points deviating by $\geq 3\sigma$
- Coherency - deviating points follow a clear trend

Detection

$\Delta\chi^2$ is fine for theoretical investigations, or formal comparison between competing models.

But as a first detection criterion, can be easily fooled in practice with real observational data.

Use is often made of some other detection criteria such as:

- Point-by-point significance - e.g 3 consecutive points deviating by $\geq 3\sigma$
- Coherency - deviating points follow a clear trend
- Confirmable - are deviating points supported by more observations

Detection

Very difficult to encode the human observer's intuition when presented with a set of real observation data.

Detection

Very difficult to encode the human observer's intuition when presented with a set of real observation data.

However Dominik & Rattenbury (2007).

Detection

Very difficult to encode the human observer's intuition when presented with a set of real observation data.

However Dominik & Rattenbury (2007).

But before we can apply detection criteria, we need to be able to fit microlensing lightcurve models!

Modelling

Modelling planetary microlensing events requires familiarity with the various lightcurve features due to planetary lens systems.

Modelling

Modelling planetary microlensing events requires familiarity with the various lightcurve features due to planetary lens systems.

Must also be aware that there is quite often more than one family of models for any given event, with similar goodness-of-fit values. May require physical reasoning to prefer one class of models over another.

Modelling

Modelling planetary microlensing events requires familiarity with the various lightcurve features due to planetary lens systems.

Must also be aware that there is quite often more than one family of models for any given event, with similar goodness-of-fit values. May require physical reasoning to prefer one class of models over another.

The position of a planet is subject to a further degeneracy: models with projected planet distance a_{p} are degenerate with those with distance $1/a_{\mathrm{p}}$.

Modelling

Point source single lens events are relatively easy to model. Standard non-linear fitting algorithms can be used to find $u_{\rm min}$, $t_E$and t_0. The blend flux parameters $F_{\rm u}$, $F_{\rm l}$ can similarly be found, either as additional parameters in the non-linear fitting routine, or through linear least-squares.

Modelling

Point source single lens events are relatively easy to model. Standard non-linear fitting algorithms can be used to find $u_{\min}$, $t_E$and t_0. The blend flux parameters F_{u}, F_{l} can similarly be found, either as additional parameters in the non-linear fitting routine, or through linear least-squares.

$$u(t) = \left[u_{\min}{}^2 + \left(\frac{(t - t_0)}{t_E} \right)^2 \right]^{\frac{1}{2}}$$

$$\mu = \frac{u^2 + 2}{u\sqrt{u^2 + 4}}$$

$$F = F_{\mathrm{l}} \cdot A(u(t)) + F_{\mathrm{u}}$$

Modelling

Finite source single lens modelling can be done using the analytic treatment of a finite source star.

- Witt & Mao, 1998

Modelling

Finite source single lens modelling can be done using the analytic treatment of a finite source star.

- Witt & Mao, 1998

Binary lens modelling with a finite source star can also be done analytically:

- Dominik, 2006

Modelling

Finite source single lens modelling can be done using the analytic treatment of a finite source star.

- Witt & Mao, 1998

Binary lens modelling with a finite source star can also be done analytically:

- Dominik, 2006

Inverse ray shooting is a numerical technique which, although slow, is robust:

- (Rattenbury, 2002)

Modelling

The software tools for generating microlensing lightcurves exist.

Modelling

The software tools for generating microlensing lightcurves exist.

There is currently no standard set of codes used by all modellers.

Modelling

The software tools for generating microlensing lightcurves exist.

There is currently no standard set of codes used by all modellers.

The lightcurve generation code used depends on the required accuracy, speed and application.

Modelling

The software tools for generating microlensing lightcurves exist.

There is currently no standard set of codes used by all modellers.

The lightcurve generation code used depends on the required accuracy, speed and application.

The fitting algorithm used also needs to be well understood.

Modelling

Apart from a general understanding of how microlensing lightcurves alter with planetary mass and position (intuition/experience), a broad coarse search over the parameter space is a good way of finding starting locations for more sophisticated fitting algorithms.

Modelling

Apart from a general understanding of how microlensing lightcurves alter with planetary mass and position (intuition/experience), a broad coarse search over the parameter space is a good way of finding starting locations for more sophisticated fitting algorithms.

The first coarse search could be a grid-wise search for example, or a random (uniform) sampling over the parameter space.

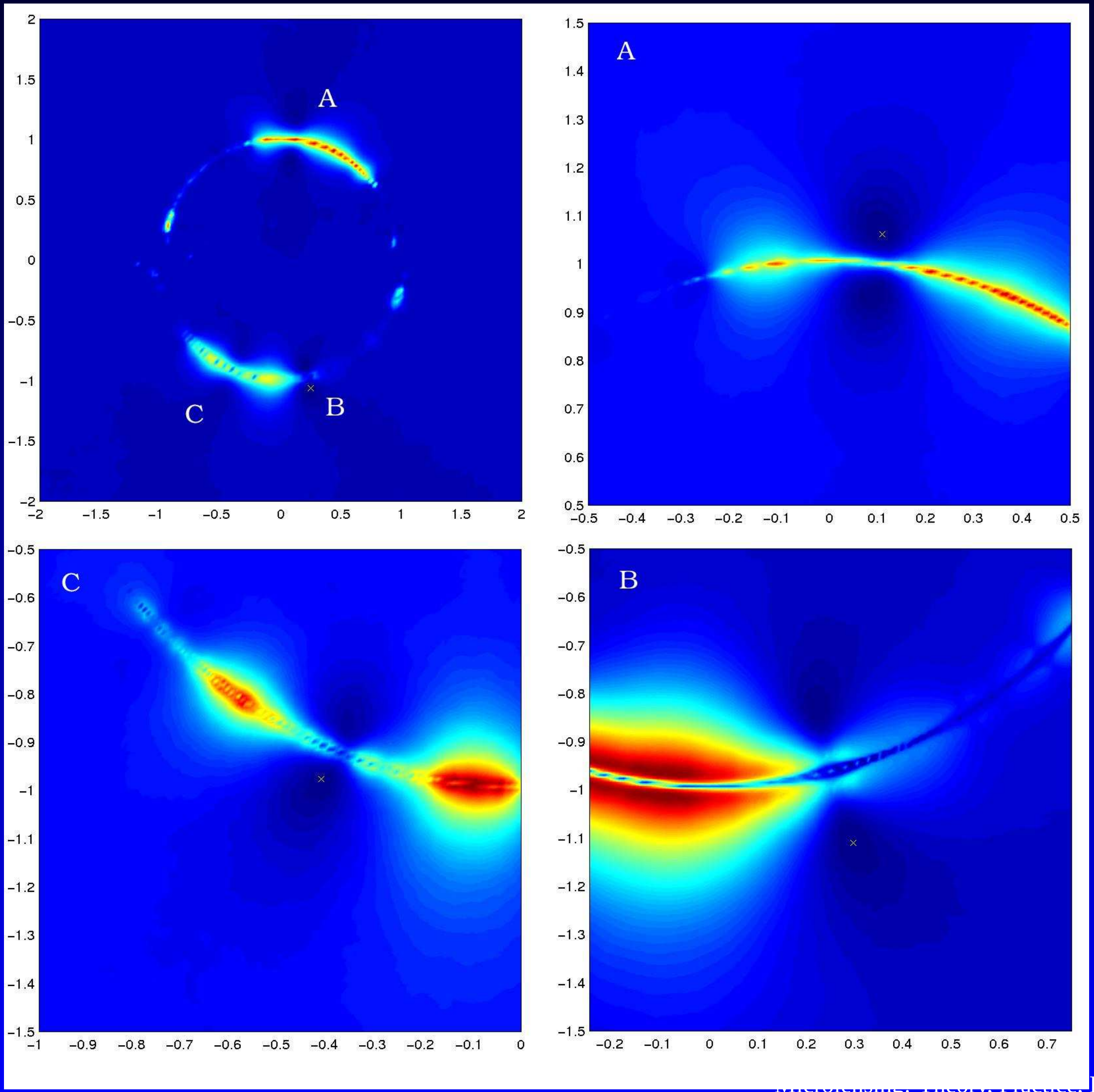
Modelling

Apart from a general understanding of how microlensing lightcurves alter with planetary mass and position (intuition/experience), a broad coarse search over the parameter space is a good way of finding starting locations for more sophisticated fitting algorithms.

The first coarse search could be a grid-wise search for example, or a random (uniform) sampling over the parameter space.

May need several restarts, but should provide some good starting locations.

Modelling



Modelling

Having determined where possible optimal parameter values might lie in the parameter space $\mathbf{\Omega} = \Omega\{u_{\rm min}\, , t_0, t_E, r_{\rm s}, x_{p_i}, y_{p_i}, q_i\}$, the next step is to apply more rigorous χ^2 minimisation procedures.

Modelling

Having determined where possible optimal parameter values might lie in the parameter space $\boldsymbol{\Omega} = \Omega\{u_{\min}, t_0, t_E, r_{\rm s}, x_{p_i}, y_{p_i}, q_i\}$, the next step is to apply more rigorous χ^2 minimisation procedures.

Simple downhill simplex non-linear algorithms (Neadler-Mead) are subject to getting trapped in local χ^2 minima.

Modelling

Having determined where possible optimal parameter values might lie in the parameter space $\mathbf{\Omega} = \Omega\{u_{\min}, t_0, t_E, r_{\rm s}, x_{p_i}, y_{p_i}, q_i\}$, the next step is to apply more rigorous χ^2 minimisation procedures.

Simple downhill simplex non-linear algorithms (Neadler-Mead) are subject to getting trapped in local χ^2 minima.

The simplex method always moving toward a lower χ^2 value. If the χ^2 manifold over the N-dimensional parameter space is not smooth, methods such as the downhill simplex can get trapped in non-optimal parameter space states.

Modelling

Having determined where possible optimal parameter values might lie in the parameter space $\boldsymbol{\Omega} = \Omega\{u_{\rm min}\,, t_0, t_E, r_{\rm s}, x_{p_i}, y_{p_i}, q_i\}$, the next step is to apply more rigorous χ^2 minimisation procedures.

Simple downhill simplex non-linear algorithms (Neadler-Mead) are subject to getting trapped in local χ^2 minima.

It is for this reason that it is suggested that the simplex method is restarted often, around the point in $\boldsymbol{\Omega}$ considered to be the set of parameter values that minimises χ^2. If the procedure recovers the same point after starting from a number of different points, the set of optimal parameters is considered secure.

Modelling

Having determined where possible optimal parameter values might lie in the parameter space $\mathbf{\Omega} = \Omega\{u_{\mathrm{min}}, t_0, t_E, r_{\mathrm{s}}, x_{p_i}, y_{p_i}, q_i\}$, the next step is to apply more rigorous χ^2 minimisation procedures.

Simple downhill simplex non-linear algorithms (Neadler-Mead) are subject to getting trapped in local χ^2 minima.

If the χ^2 manifold over $\mathbf{\Omega}$ is sufficiently complicated, there is a greater risk that the results returned from the simplex methods are not global minima.

Modelling

The Metropolis-Hastings Markov Chain Monte Carlo (MHMCMC) algorithm (Grey 1992) can be used to minimise a function over a parameter space in much the same way as the simplex method described above.

Modelling

The Metropolis-Hastings Markov Chain Monte Carlo (MHMCMC) algorithm (Grey 1992) can be used to minimise a function over a parameter space in much the same way as the simplex method described above.

The main important difference is that each next point in Ω is chosen with a certain probability. This is in contrast to the simplex method, which always accepts a better candidate state compared to the current state.

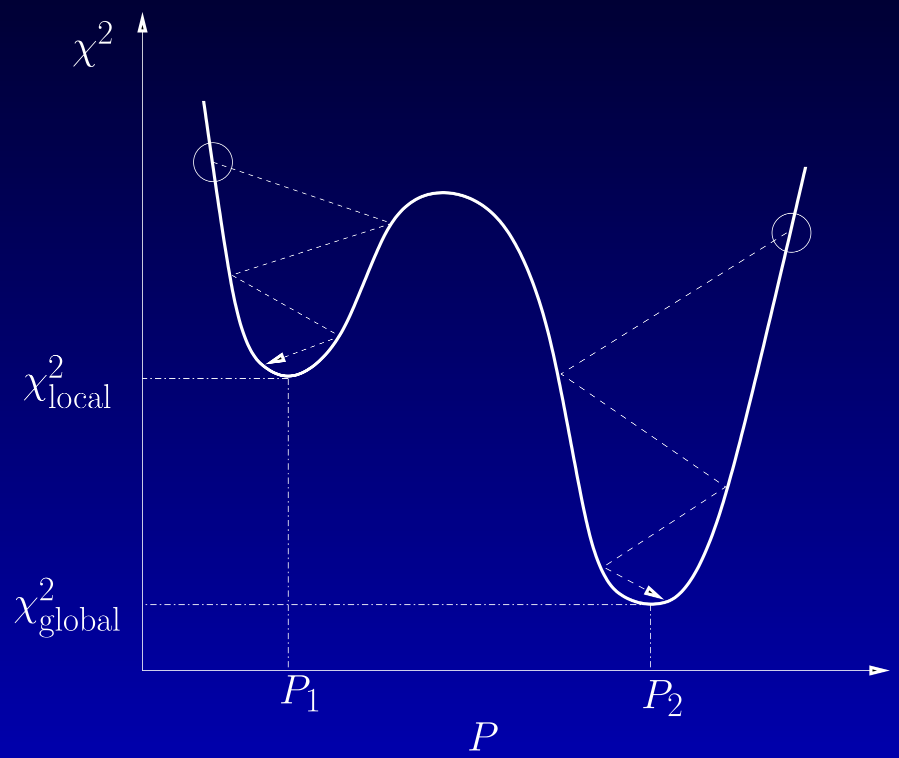
Modelling

The Metropolis-Hastings Markov Chain Monte Carlo (MHMCMC) algorithm (Grey 1992) can be used to minimise a function over a parameter space in much the same way as the simplex method described above.

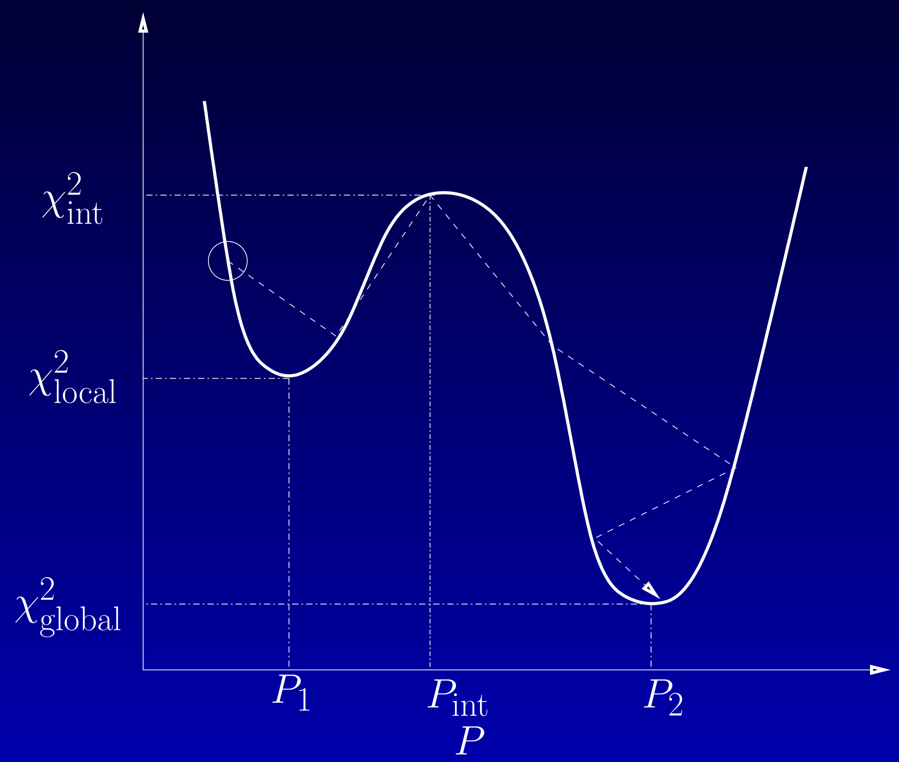
The main important difference is that each next point in Ω is chosen with a certain probability. This is in contrast to the simplex method, which always accepts a better candidate state compared to the current state.

This feature of the MHMCMC algorithm allows the method to occasionally disregard a (possibly better) candidate state. A chance therefore exists for the algorithm to find a more favourable region on the χ^2 manifold.

Modelling



χ^2
χ^2_{local}
χ^2_{global}
P_1
P_2
P



χ^2_{int}
χ^2_{local}
χ^2_{global}
P_1
P_{int}
P_2
P

Modelling

Assume that the current set of parameters is Ω.

Modelling

Assume that the current set of parameters is $\mathbf{\Omega}$.

Let $\{Z_k\}_{k=1}^{M}$ define a Markov chain, length M. Each element of Z_k is a set of parameters: $Z_k = \mathbf{\Omega}_k$.

Modelling

The next element in the chain, Z_{k+1} is determined as follows:

1. Choose a candidate set of parameters, **Ω'**, by varying one or more of the parameters:

$$
\begin{aligned}
{u_{\min}}' &= u_{\min} + \mathcal{R}_{u_{\min}} \qquad (0)\\
t_0' &= t_0 + \mathcal{R}_{t_0}\\
t_E' &= t_E + \mathcal{R}_{t_E}\\
\vdots &= \vdots
\end{aligned}
$$

where $\mathcal{R}$ is an appropriately scaled random number.

Modelling

2. The candidate state, $\mathbf{\Omega}$', is accepted with probability:

$$\alpha(\mathbf{\Omega}'|\mathbf{\Omega}) = \min\left\{1, \mathcal{F}(f'(t), f(t))\right\}$$

where $f(t)$ is the model function calculated using the current parameter set, $\mathbf{\Omega}$. $f'(t)$ is the model function, calculated using the candidate parameter set, $\mathbf{\Omega}$'. $\mathcal{F}$ is a function based on the difference between the current and candidate models. If the candidate state $\mathbf{\Omega}$' is not accepted, set $Z_{k+1} = \mathbf{\Omega}_k$.

Modelling

$\alpha(\mathbf{\Omega}'|\mathbf{\Omega})$ is the acceptance ratio. The function $\mathcal{F}$ essentially compares the χ^2 values arising from fitting the light curve generated using the candidate parameters **Ω**', to that of the current set.

$$\mathcal{F} = \exp\left[\frac{\chi^2 - \chi\prime^2}{2}\right]$$

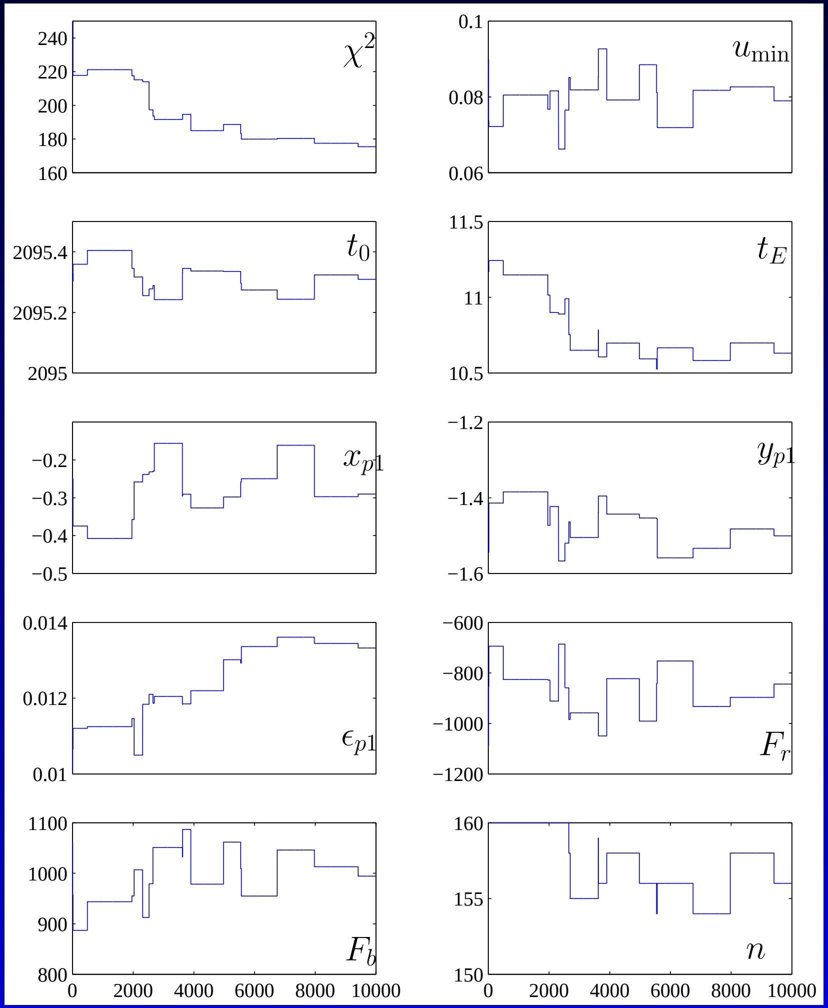
Modelling

$\alpha(\mathbf{\Omega}'|\mathbf{\Omega})$ is the acceptance ratio. The function $\mathcal{F}$ essentially compares the χ^2 values arising from fitting the light curve generated using the candidate parameters **Ω**', to that of the current set.

$$\mathcal{F} = \exp\left[\frac{\chi^2 - \chi\prime^2}{2}\right]$$

The above two procedures are iterated M times. Each parameter is allowed to vary, and the algorithm settles at a point in the parameter space which minimises χ^2.

Modelling



χ^2
240
220
200
180
160
$u_{\rm min}$
0.1
0.08
0.06
t_0
2095.4
2095.2
2095
t_E
11.5
11
10.5
x_{p1}
−0.2
−0.3
−0.4
−0.5
y_{p1}
−1.2
−1.4
−1.6
ϵ_{p1}
0.014
0.012
0.01
F_r
−600
−800
−1000
−1200
F_b
1100
1000
900
800
n
160
155
150
0 2000 4000 6000 8000 10000
0 2000 4000 6000 8000 10000

Modelling

We can also impose an annealing schedule, $T(n)$:

$$\mathcal{F} = \exp\left[\frac{\chi^2 - \chi'^2}{2T}\right]$$

If $T \rightarrow 0$, the distribution will converge on the optimal modal space in the state space. Given a distribution in equilibrium, the “temperature”, T, can slowly be lowered and allow the distribution to find its new equilibrium. As $T \rightarrow 0$ the equilibrium will settle on the mode state(s).

Modelling

	$\Delta\chi^2 > 0$		$\Delta\chi^2 < 0$	
T	$\alpha(z'\|z)$	Accept z'	$\alpha(z'\|z)$	Accept z'
> 1	1	always	≤ 1	often
1.0	1	always	$0 < \alpha < 1$	sometimes
$0 < T < 1$	1	always	$0 < \alpha \ll 1$	seldom
0.0	1	always	0	never

$\Delta\chi^2 = \chi^2 - \chi'^2$, where χ'^2 corresponds to the candidate state, and χ^2 to the current state.

Modelling

- MHMCMC is excellent for fitting microlensing lightcurves
- MHMCMC fitting can be optimised in many ways.
- MHMCMC can be parallized easily.
- There is a lot of information in the equilibrium state of each parameter.

Genetic algorithms such as Particle Swarm Optimisation, Artificial Neural Networks and Self Organising Maps are other possible categorization methods.