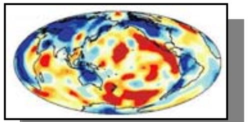
$$\sigma(d,m) = k\frac{\rho(d,m)\theta(d,m)}{\mu(d,m)}$$

Inverse problems: Special solutions



Albert Tarantola

- The resolution of inverse problems

- Special cases

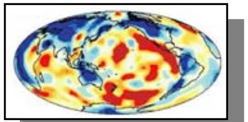
- independent errors
- negligible modeling errors
- negligible observational errors
- Gaussian hypothesis
- linear forward problem

- Our real example: hypocentre location

This lecture follows **Tarantola**, Inverse problem theory, p. 1-88.

$$\sigma(d,m) = k\frac{\rho(d,m)\theta(d,m)}{\mu(d,m)}$$

Resolution of inverse problems



The final information on the model space is given by the marginal a posteriori pdf

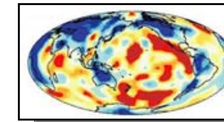
$$\sigma_M(\mathbf{m}) = \int_D d\mathbf{d}\,\sigma(\mathbf{d},\mathbf{m}) = \rho_M(\mathbf{m})\int_D d\mathbf{d}\,\frac{\rho_D(d)\theta(\mathbf{d}\,|\,\mathbf{m})}{\mu_D(\mathbf{d})}$$

This really is **the** solution to the inverse problem. One can use σ(m)
To obtain any kind of information one wishes, for example

- mean values
- maximum likelihood values
- error bars

.. etc.

Special cases (formally)



1. Negligible modelisation errors (compared to other errors)

$$\Theta(\mathbf{d}\,|\,\mathbf{m}) = \delta(\mathbf{d} - \mathbf{g}(\mathbf{m}))$$

$$\mathbf{d} = \mathbf{g}(\mathbf{m})$$

d=g(m) is the exact solution of the forward problem. For the
A posteriori pdf in the model space we obtain:

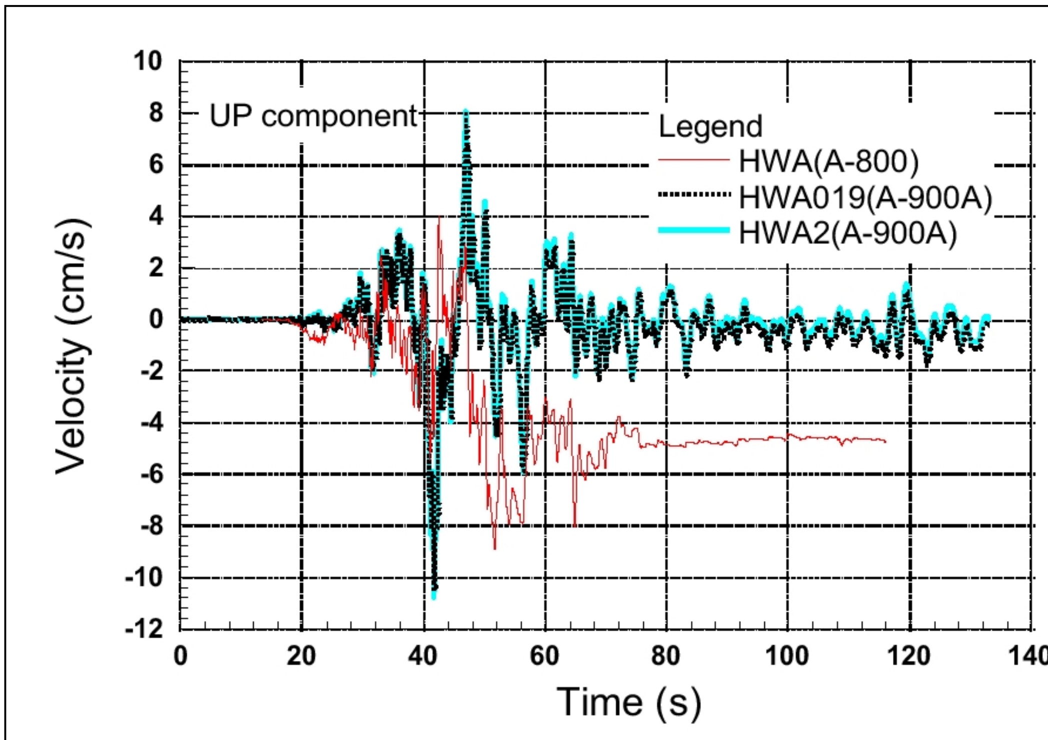
$$\boxed{\sigma_M(\mathbf{m}) = \rho_M(\mathbf{m})\frac{\rho_D(\mathbf{d})}{\mu_D(\mathbf{d})}\Big|_{\mathbf{d}=\mathbf{g}(\mathbf{m})}}$$

In words: We are sure of the correct physical description and the (e.g. analytical, numerical, ...) solution to the problem. The only alteration of the solution to the inverse problem comes from Possible errors in the data $\rho_M(d)$ and the a priori information on the model parameters $\rho_M(m)$.

$$\sigma(d,m) = k\frac{\rho(d,m)\theta(d,m)}{\mu(d,m)}$$

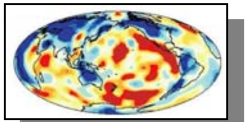
Errors: example

Velocity seismograms derived from three accelerometers at the same location recording the 1999 Chi-Chi earthquake in Taiwan, they should really be the same …



$$\sigma(d,m) = k\frac{\rho(d,m)\theta(d,m)}{\mu(d,m)}$$

Special cases (formally)



2. Negligible observational errors (compared to other errors)

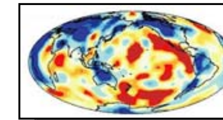
$$\rho_D(\mathbf{d}) = \delta_M(\mathbf{d} - \mathbf{d_{obs}})$$

.. we are absolutely certain about the quality of our data, then we obtain …

$$\boxed{\sigma_M(\mathbf{m}) = \rho_M(\mathbf{m})\frac{\theta(\mathbf{d}_{obs}\,|\,\mathbf{m})}{\mu_D(\mathbf{d}_{obs})}}$$

In words: Although we are sure about the data we allow for uncertainties in our forward modelling $\Theta(.)$. On top of this the a priori $\rho(m)$ information influences the solution to our problem.

Special cases (formally)



3. Gaussian modeling and observational errors

$$\Theta(\mathbf{d}\,|\,\mathbf{m}) \propto \exp\left\{-\frac{1}{2}(\mathbf{d}-\mathbf{g(m)})^t C_T^{-1}(\mathbf{d}-\mathbf{g(m)})\right\}$$

$$\rho_D(\mathbf{m}) \propto \exp\left\{-\frac{1}{2}(\mathbf{d}-\mathbf{d_{obs}})^t C_D^{-1}(\mathbf{d}-\mathbf{d_{obs}})\right\}$$

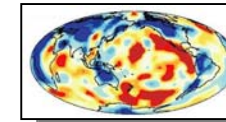
... with the final solution (a posteriori pdf)

$$\boxed{\sigma_M(\mathbf{m}) \propto \rho(\mathbf{m})\exp\left\{-\frac{1}{2}(\mathbf{d}_{obs}-\mathbf{g(m)})^t C^{-1}(\mathbf{d}_{obs}-\mathbf{g(m)})\right\}}$$

... note here that $C=C_D+C_T$

In words: In the Gaussian assumption, observational errors and modeling errors simply combine by addition of the respective covariance operators (even for nonlinear forward problems).

More on Gauss!



The Gaussian hypothesis (least-squares criterion)

1777-1855

… let us now also assume that the a priori information $\rho(m)$ is Gaussian, we then obtain the most commonly used solution

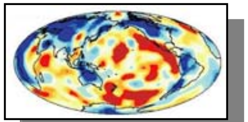
$$\sigma_M(\mathbf{m}) \propto$$

$$\exp\left\{-\frac{1}{2}\left((\mathbf{d}_{obs} - \mathbf{g}(\mathbf{m}))^t C_D^{-1}(\mathbf{d}_{obs} - \mathbf{g}(\mathbf{m})) + (\mathbf{m} - \mathbf{m}_{prior})^t C_M^{-1}(\mathbf{m} - \mathbf{m}_{prior})\right)\right\}$$

… Even though we are interested in the general nonlinear case $d=g(m)$ let us quickly explore the linear case when $d=g(m)=Gm$, G being a matrix operator.

$$\sigma(d,m) = k\frac{\rho(d,m)\theta(d,m)}{\mu(d,m)}$$

Gauss- the linear case!



The Gaussian hypothesis (least-squares criterion)

1777-1855

... the linear case leads to the following

$$\mathbf{d} = \mathbf{G}\mathbf{m}$$

$$\sigma_M(\mathbf{m}) \propto \exp\{-S(\mathbf{m})\}$$

$$S(\mathbf{m}) =$$

$$\exp\left\{-\frac{1}{2}\left((\mathbf{Gm}-\mathbf{d}_{\mathbf{obs}})^t C_D^{-1}(\mathbf{Gm}-\mathbf{d}_{\mathbf{obs}}) + (\mathbf{m}-\mathbf{m}_{prior})^t C_M^{-1}(\mathbf{m}-\mathbf{m}_{prior})\right)\right\}$$

by differentiation we can find the mean model <m> of this function

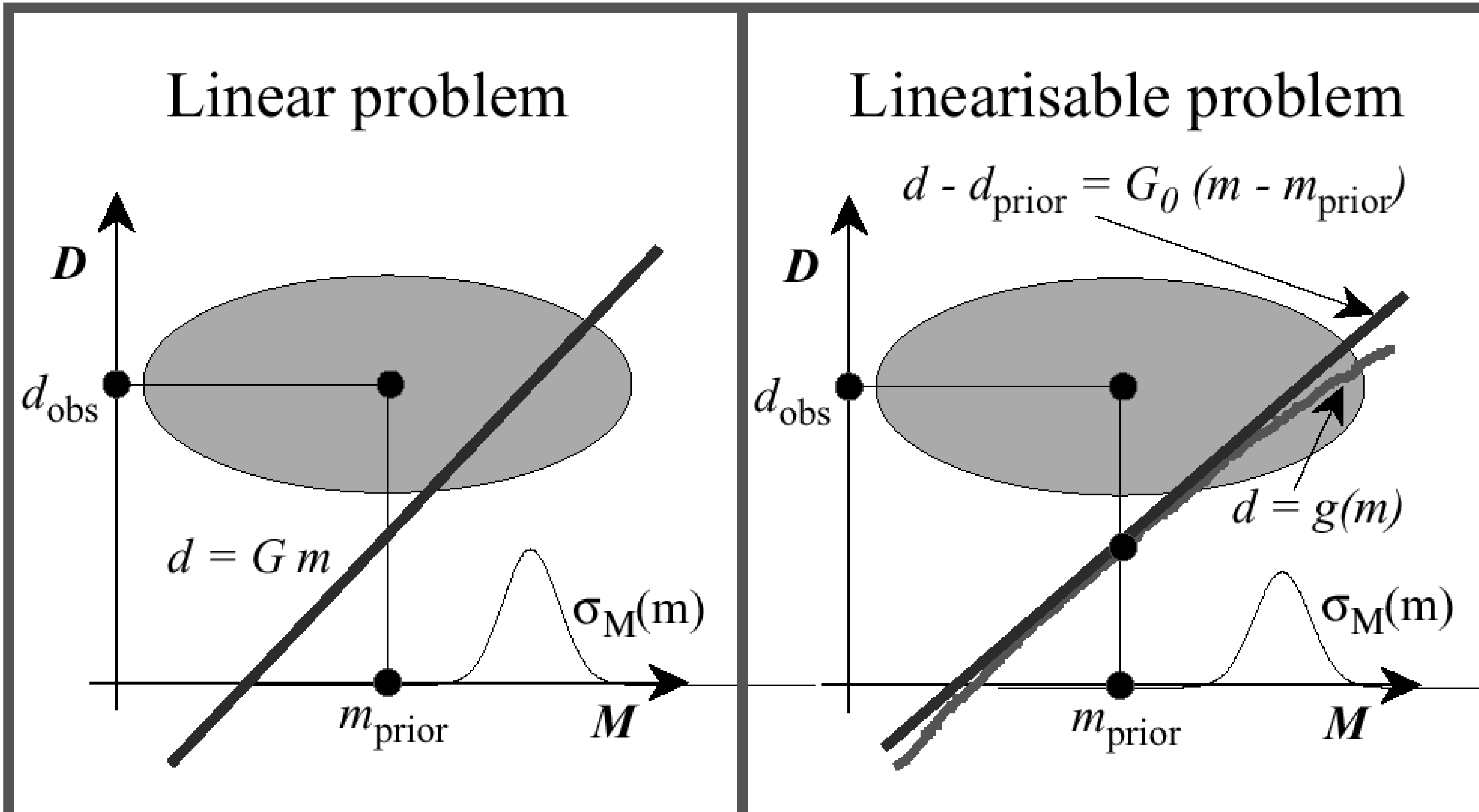
$$<\mathbf{m}> = \left(\mathbf{G}^t\mathbf{C}_\mathbf{D}^{-1}\mathbf{G} + \mathbf{C}_\mathbf{M}^{-1}\right)^{-1}\left(\mathbf{G}^t\mathbf{C}_\mathbf{D}^{-1}\mathbf{d}_{\mathbf{obs}} + \mathbf{C}_\mathbf{M}^{-1}\mathbf{m}_{\mathbf{prior}}\right)$$

This model is the maximum likely model as it has the highest probability.
The a posteriori probability density in the model space is Gaussian!

$$\sigma(d,m) = k\,\frac{\rho(d,m)\,\theta(d,m)}{\mu(d,m)}$$

Gauss - graphically!

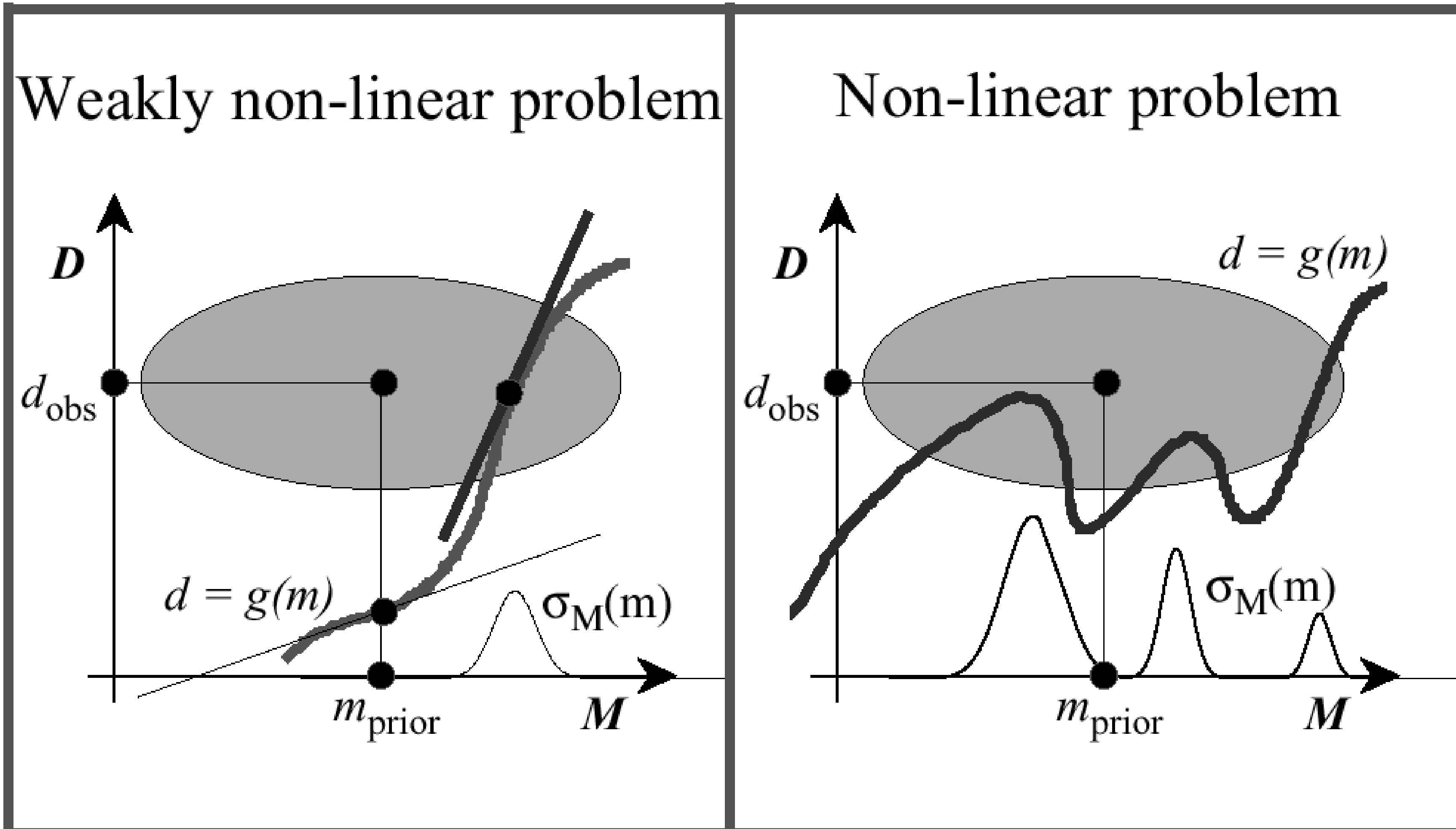
... here we assume exact theoretical data d



$$\sigma(d,m) = k\frac{\rho(d,m)\theta(d,m)}{\mu(d,m)}$$

Gauss - graphically!

... and it s getting wilder ...

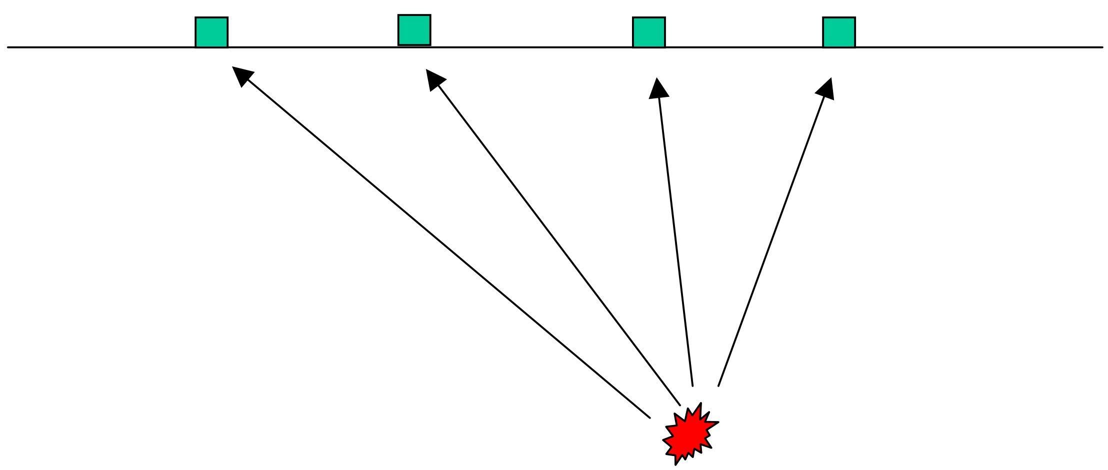


$$\sigma(d,m) = k\frac{\rho(d,m)\theta(d,m)}{\mu(d,m)}$$

Probabilistic estimation of hypocenter

The problem: The seismic waves recorded at a seismic network carry Information on the location of an earthquake. The parameters For this problem are simply the x,y,z coordinates of Euclidean space. However, we are not interested in one particular solution but in investigation the amount of informatino we have on the space of possible models in x,y,z! To further simplify we consider The problem in the x,z plane only.

Seismic network



$$\sigma(d,m) = k\,\frac{\rho(d,m)\theta(d,m)}{\mu(d,m)}$$

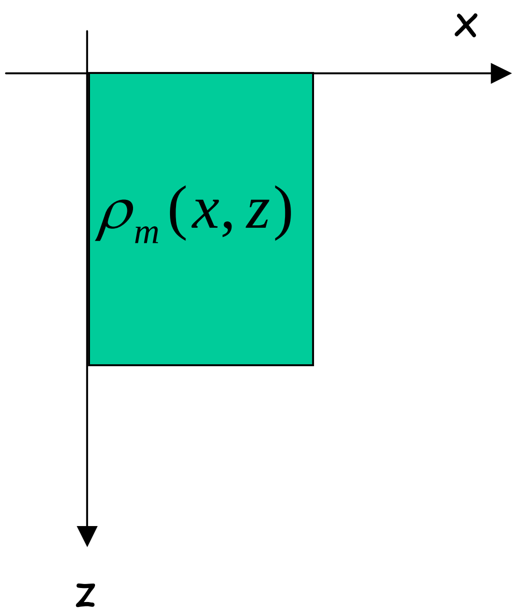
Hypocenter location

A priori information:

The unknowns (model parameters) are the hypocentral coordinates x,z and origin time T. We assume a priori information on our parameters.

$$\rho_m(x,z,T)$$ prior information

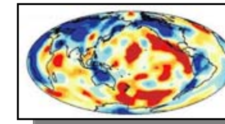
$$\mu_m(x,z,T) = k$$ non-informative prior



We assume that the prior information is constant in side the region 0<x<60km and 0<z<50km. The prior information on t is constant (non-informative).

$$\sigma(d,m) = k\frac{\rho(d,m)\theta(d,m)}{\mu(d,m)}$$

Hypocenter location

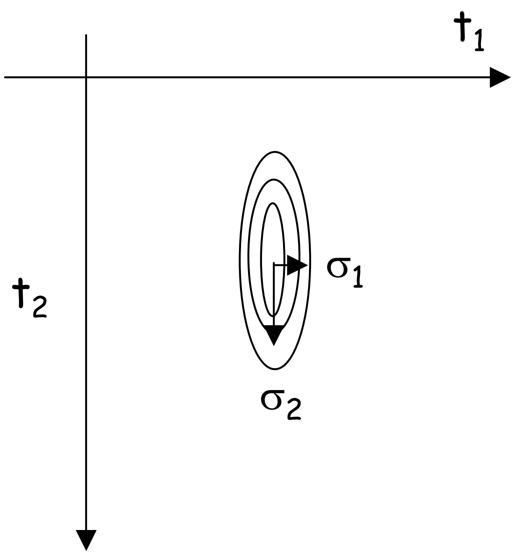


A priori information (data):

We have four arrival times of the earthquake $\{t_1, t_2, t_3, t_4\}$ at four observatories at locations $\{x_i, y_i\}$. The measurements are described by the probability

$$\rho_d(t_1, t_2, t_3, t_4)$$ prior information (data)

$$\mu_d(t_1, t_2, t_3, t_4) = k$$ non-informative prior (data)

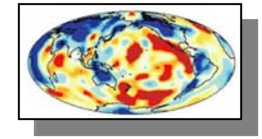


We will assume Gaussian errors for each t_i with varian σ_i:

$$\rho_d(t_1, t_2, t_3, t_4) = k \exp\left(-\frac{1}{2}\frac{(t_1 - t_1^{obs})^2}{\sigma_1^2}\right) \exp\left(-\frac{1}{2}\frac{(t_2 - t_2^{obs})^2}{\sigma_2^2}\right)$$
$$\exp\left(-\frac{1}{2}\frac{(t_3 - t_3^{obs})^2}{\sigma_3^2}\right) \exp\left(-\frac{1}{2}\frac{(t_4 - t_4^{obs})^2}{\sigma_4^2}\right)$$

$$\sigma(d,m) = k\frac{\rho(d,m)\theta(d,m)}{\mu(d,m)}$$

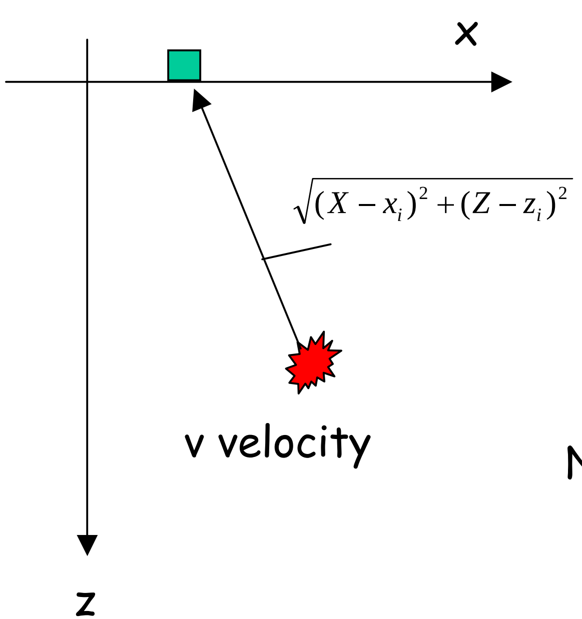
Hypocenter location: forward problem



To solve the forward problem d=g(m) we need to calculate the arrival times t_i at the seismic stations x_i,z_i for the hypocentral coordinates X,Z

$$t_i = f(X,Z,T)$$

$$t_i^{cal} = T + \frac{\sqrt{(X-x_i)^2 + (Z-z_i)^2}}{v}$$



Here we assume a constant velocity model.
Is this a good approximation?
We also assume that our modeling is exact.

Now we can formulate the solution to our inverse problem!

Hypocenter location: solution

We can now combine the a priori information with that obtained from our experiment:

$$\sigma_m(X,Z,T) = k\rho_m(X,Z,T)\rho_d(t_1,t_2,t_3,t_4)\Big|_{t_i=f(X,Z,T)}$$

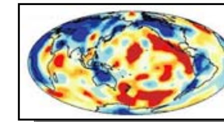
Now let us calculate this a posteriori probability density function.

The stations: $s_1(5,0)$, $s_1(10,0)$, $s_1(15,0)$, $s_1(25,0)$ in km
The velocity model: v=5 km/s

The data: t_{obs}= (30.3, 29.4, 28.6, 28.3)
σ = (0.1, 0.2, 0.1, 0.2)

$$\sigma(d,m) = k\,\frac{\rho(d,m)\theta(d,m)}{\mu(d,m)}$$

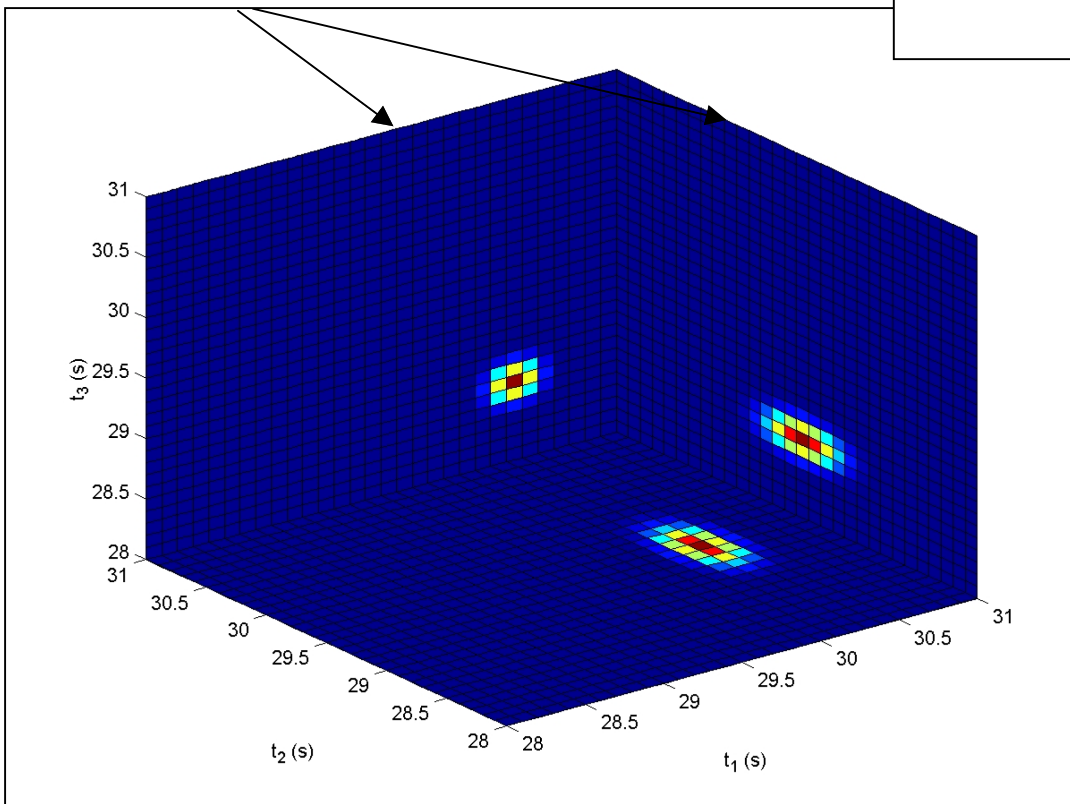
Hypocenter: data uncertainties



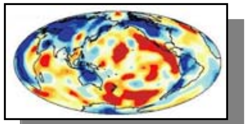
The stations: $s_1(5,0)$, $s_1(10,0)$, $s_1(15,0)$, $s_1(25,0)$ in km
The velocity model: v=5 km/s

The data: t_{obs}= (30.3, 29.4, 28.6, 28.3)
σ = (0.1, 0.2, 0.1, 0.2)

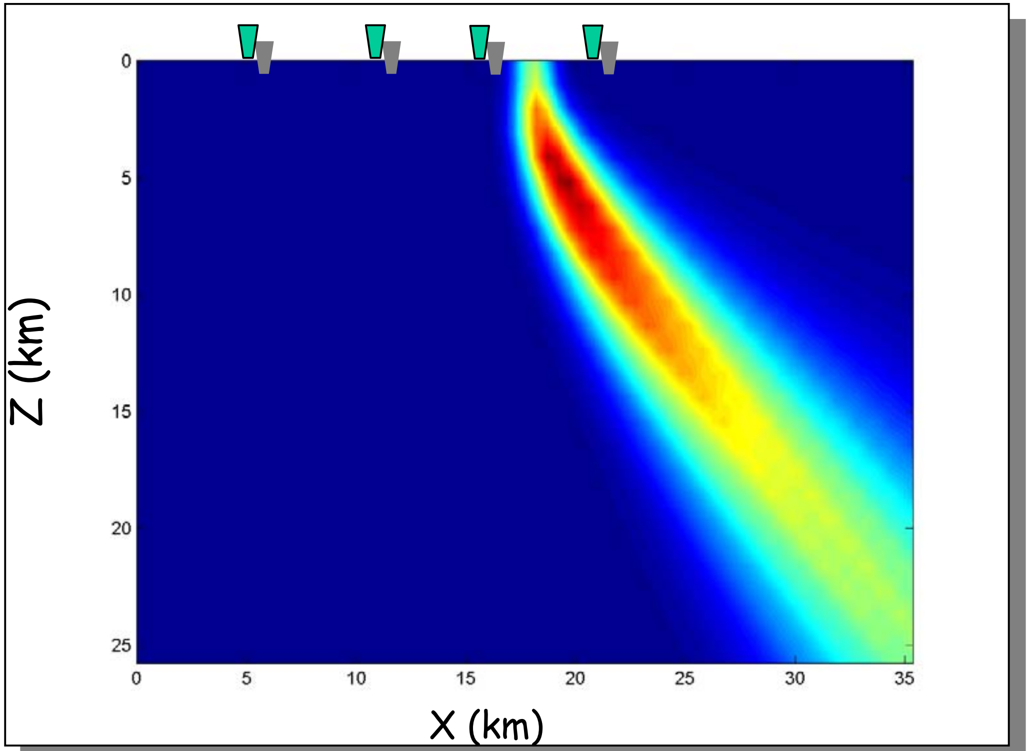
Marginal prior probabilities



$$\sigma(d,m) = k\frac{\rho(d,m)\theta(d,m)}{\mu(d,m)}$$



The solution to our inverse problem (x,z – space)

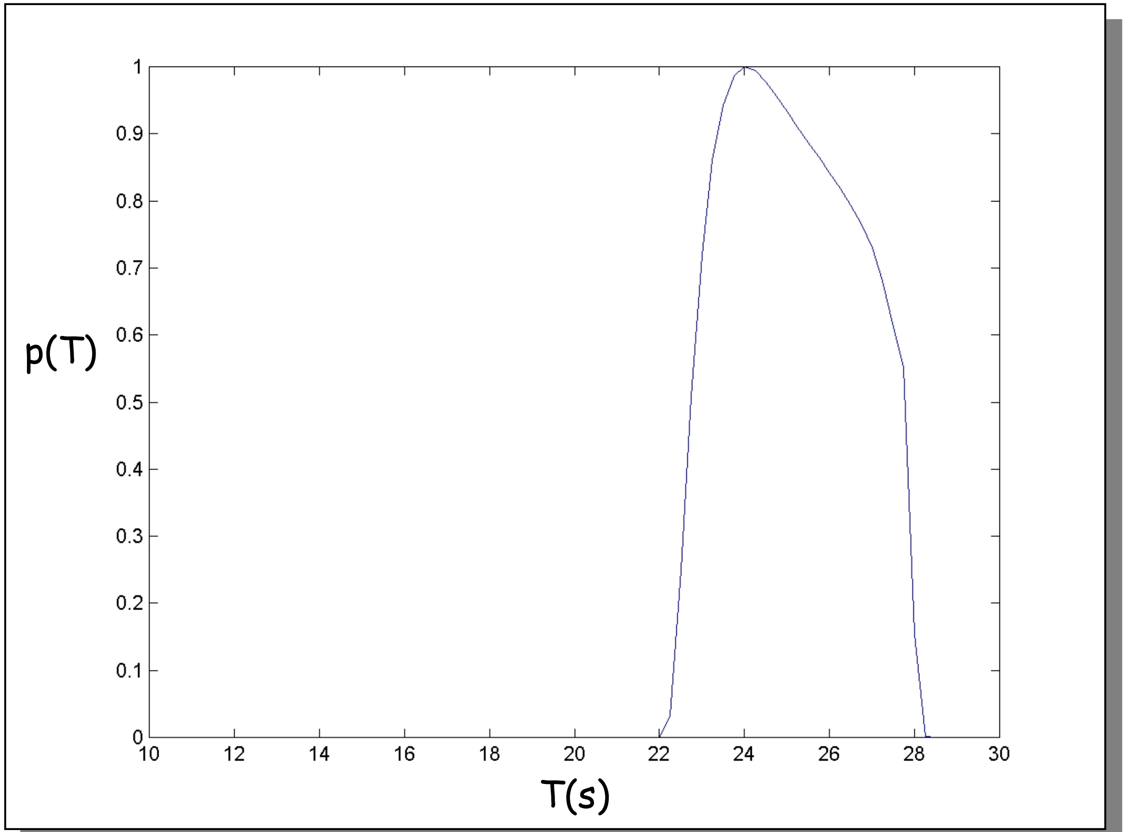


This is the marginal probability density in the x,z space. One can Now calculate any desired values (maximum likelihood point, standard Deviations, etc.)

$$\sigma(d,m) = k\frac{\rho(d,m)\theta(d,m)}{\mu(d,m)}$$

Hypocenter location: a posteriori

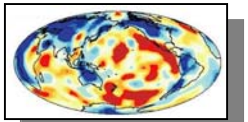
The solution to our inverse problem (T- space)



This graph shows the uncertainties on the origin time of the earthquake

$$\sigma(d,m)=k\frac{\rho(d,m)\theta(d,m)}{\mu(d,m)}$$

Summary



The hypocenter problem has shown us most benefits and difficulties of probabilistic inverse theory

Benefits:

The visual representation of the marginal probabilities seems to be an optimal way of describing information we have on the hypocenter!

Difficulties:

We paid a high price: calculating lots and lots of models with extremely low probability. We sampled the a posteriori probability going through the whole model space without thought.

There must be better ways of doing this!