Threshold Effects in Parameter Estimation as Phase Transitions in Statistical Physics

Neri Merhav

Department of Electrical Engineering

Technion—Israel Institute of Technology

Haifa 32000, Israel

ITW 2011, Paraty, Brazil, October 2011.

The Threshold Effect

Consider the model

$$r(t) = s_m(t) + n(t), \qquad -T/2 \leq t < T/2,$$

where:

$s_m(t) =$ a waveform parameterized by m;

$n(t) =$ AWGN with spectral density $N_0/2$.

Conveying information via a parameter m by modulating it in $s_m(t)$:

Shannon–Kotel'nikov mappings (Floor '08, Floor & Ramstad '09, Hekland '07, Ramstad '02 + references).

Nonlinear modulation $\Longrightarrow$ threshold effet:

Below some critical SNR, anomalous errors dominate the MSE.

The Threshold Effect (Cont'd)

- Not an artifact of a particular modulation/estimation scheme: it cannot be avoided.
- In the wideband regime, the threshold effect becomes abrupt: Pr$\{$anomaly$\}$ jumps from ~ 0 to ~ 1.

In this talk, we relate the abrupt threshold effect to phase transitions in statistical physics.

Motivation: the physical perspective gains our insight on the problem.

Some Background (W & J, ‘65)

In the simple case of a linear model

$$r(t) = m \cdot s(t) + n(t), \qquad -T/2 \leq t < T/2$$

the ML estimator always achieves

$$\text{MSE} = \text{CRLB} = \frac{N_0}{2E},$$

where E is the energy of $\{s(t)\}$: $\Leftrightarrow$ No threshold effect.

The only way to improve: non–linear modulation of $s_m(t)$.

Background (Cont'd) – Nonlinear Modulation

Now MSE depends not only on E: Let

$$s_m(t) \approx s_{m_0}(t) + (m - m_0) \cdot \dot{s}_{m_0}(t).$$

like the linear case with $\dot{s}_{m_0}(t)$ in the role of $s(t)$. Thus, at high SNR,

$$\text{MSE} \approx \text{CRLB} \approx \frac{N_0}{2\dot{E}},$$

where $\dot{E} =$ energy of $\dot{s}_{m_0}(t)$, which depends on more details.

For example, if $s_m(t) = s(t - m)$, $\dot{E} = W^2 E$, where

$$W = \sqrt{\frac{1}{E} \int_{-\infty}^{\infty} \mathrm{d}f \cdot (2\pi f)^2 S(f)} \qquad \text{Gabor bandwidth}$$

Why not increase W without a limit?

Background (Cont'd) – Geometry of Anomalous Errors

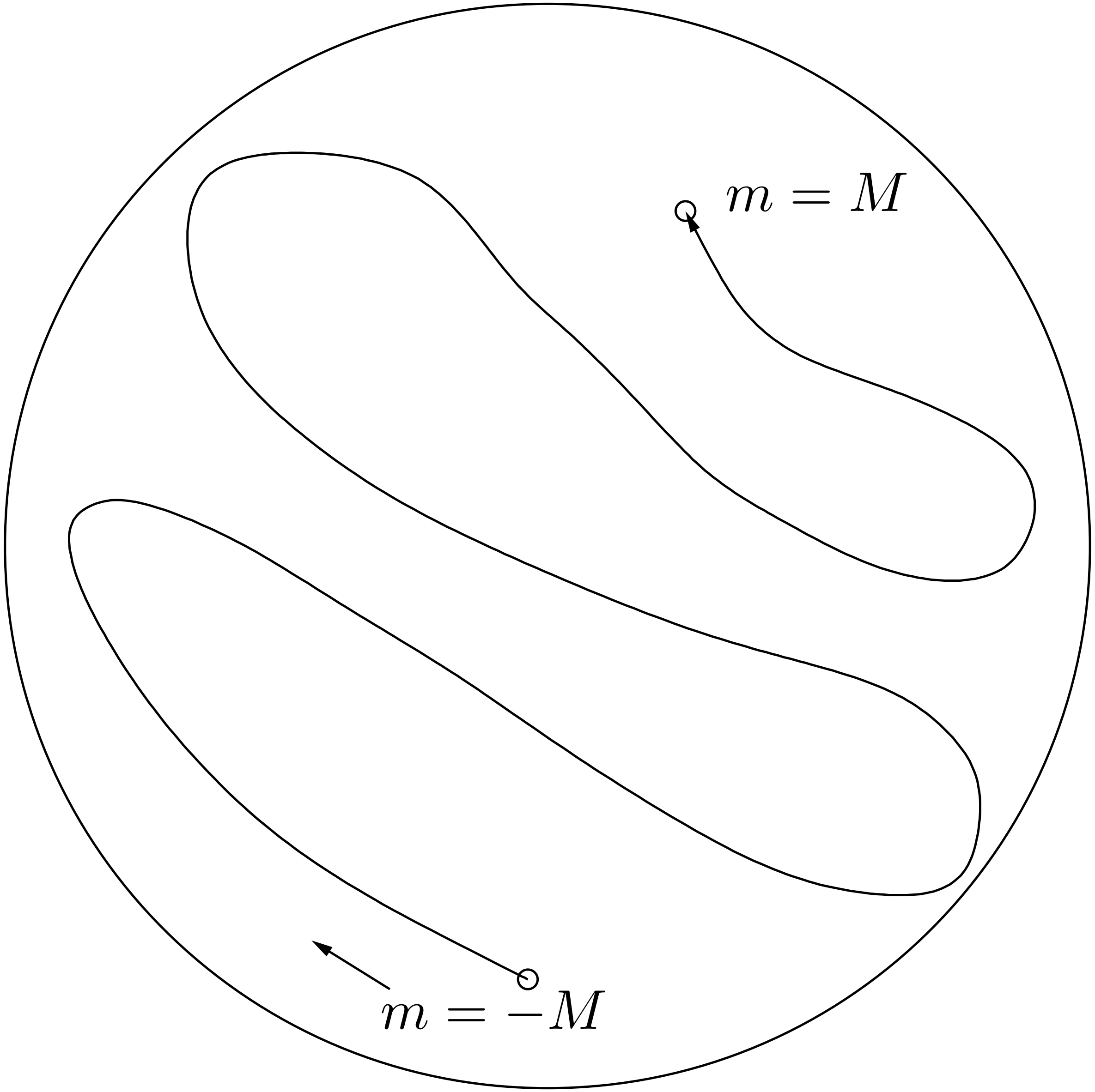
Let $\bar{s}(m) = (s_1(m), \ldots, s_K(m))$ = representation of $s_m(t)$ by K orthonormal basis functions and consider the locus of $\{\bar{s}(m), \ -M \leq m \leq M\}$ in $\mathbb{R}^K$.

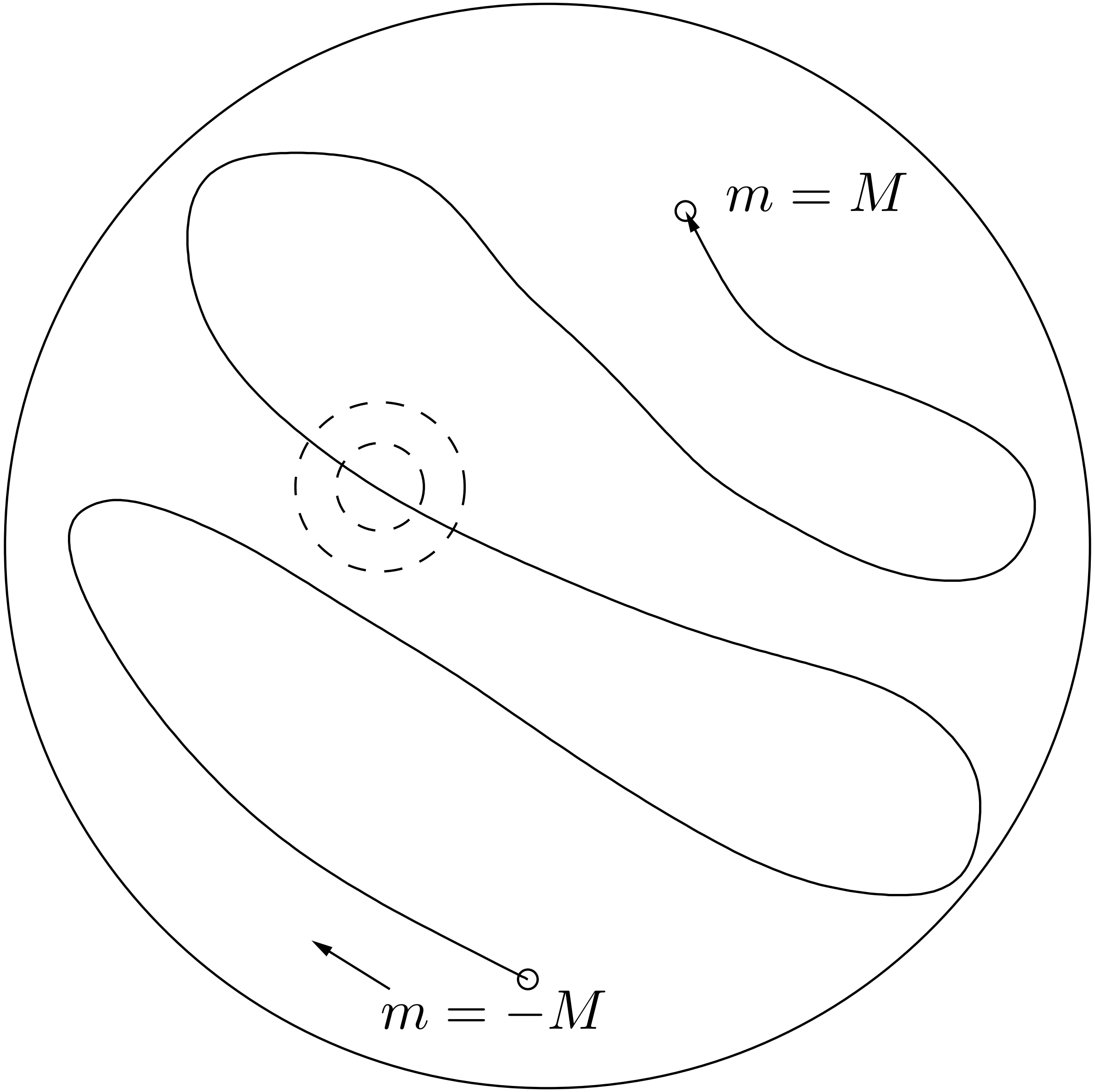
Assuming that E is independent of m, the locus lies on the hypersurface of the K–dimensional sphere of radius $\sqrt{E}$.

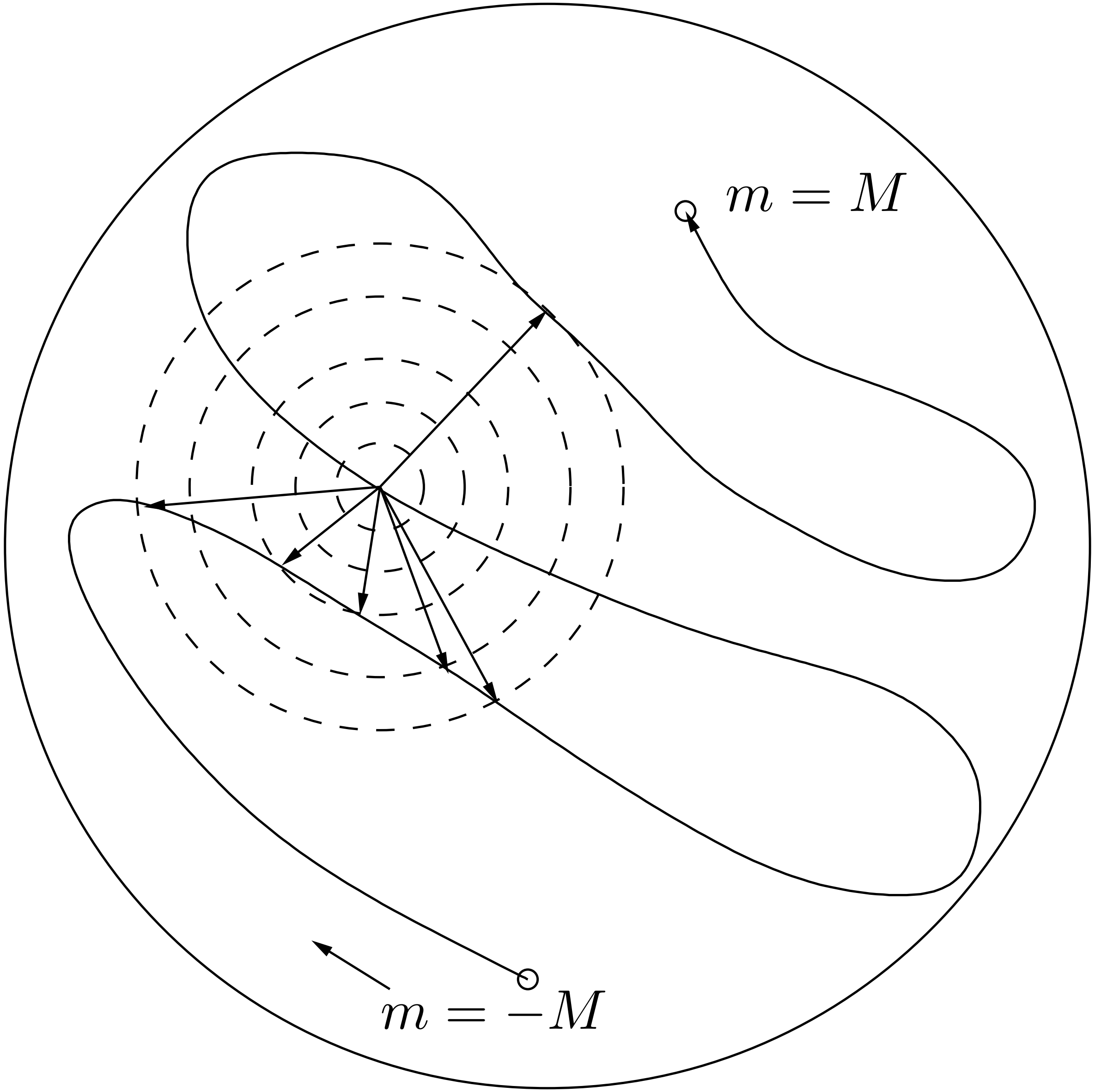
The length of the curve

$$L = \int_{-M}^{M} \mathrm{d}m \sqrt{\sum_i \dot{s}_i^2(m)} = 2M\sqrt{\dot{E}}.$$

High–SNR MSE $\downarrow$ with $\dot{E}$, we want $\dot{E} \uparrow$, thus $L \uparrow$.


$m = M$
$m = -M$


$m = M$
$m = -M$



$m = M$
$m = -M$

Anomalous Errors (Cont'd)

L – limited by the need of safe distances between folds – hot dog packing.
Maximum achieable $L \sim e^{CT}$, $C = P/N_0$ (PPM).
For PPM, $K \sim 2WT$,

$$\text{MSE} \approx \underbrace{\frac{N_0}{2W^2E}}_{\text{small error}} + \underbrace{4M^2 \cdot 2WT \cdot e^{-E/(2N_0)}}_{\text{anomalous error}}$$

For fixed W, anomalous error term $\uparrow$ gracefully as $E/N_0 \downarrow$.
For a better balance between terms – let $W \sim e^{RT}$.

$$\text{MSE} \approx \frac{N_0}{2E} e^{-2RT} + 4M^2 \cdot e^{-TE(R)} \qquad R < C$$

where $E(R) =$ reliability function of AWGN channel.
For $W \sim e^{RT}$, anomalous error $\uparrow$ abruptly as $E/N_0 \downarrow$, like a phase transition.

Purpose of this work: study the abrupt threshold effect from the viewpoint of the physics of phase transitions.

A Physical Perspective on the Threshold Effect

Consider the PPM model

$$r(t) = s(t - mT) + n(t), \quad |t| \leq T/2, \quad |m| \leq M, \quad M < \frac{1}{2}.$$

Imagine that $m \sim U[-M/2, +M/2]$, then

$$\begin{aligned} P(m|\mathcal{R}) &= \frac{P(m)P(\mathcal{R}|m)}{\int \mathrm{d}m' P(m')P(\mathcal{R}|m')} \\ &= \frac{\exp\left\{\frac{2}{N_0}\int_{-T/2}^{T/2} r(t)s(t-mT)\mathrm{d}t\right\}}{\int_{-M}^{M} \mathrm{d}m' \exp\left\{\frac{2}{N_0}\int_{-T/2}^{T/2} r(t)s(t-m'T)\mathrm{d}t\right\}} \end{aligned}$$

where $\mathcal{R} = \{r(t), \ |t| \leq T/2\}$. This can be viewed as the Boltzmann distribution with inverse temperature $\beta = 2/N_0$ and Hamiltonian (energy function)

$$\mathcal{E}(m) = -\int_{-T/2}^{T/2} r(t)s(t - mT)\mathrm{d}t.$$

A Physical Perspective ... (Cont'd)

Borrowing from the concept of finite–temperatue decoding [Ruján '93], define

$$P_\beta(m|\mathcal{R}) = \frac{\exp\left\{\beta \int_{-T/2}^{T/2} r(t)s(t-mT)\mathsf{d}t\right\}}{\int_{-M}^{M} \mathsf{d}m' \exp\left\{\beta \int_{-T/2}^{T/2} r(t)s(t-m'T)\mathsf{d}t\right\}}.$$

Motivation: a degree of freedom in case of uncertainty; simulated annealing, analysis of ML estimation.

Meaningful choices of β:
$\beta = 0$ – prior; $\beta = 2/N_0$ – natural; $\beta \to \infty$ – ML estimator dominates.

Define a partition function:

$$\zeta(\beta) = \int_{-M}^{M} \mathsf{d}m \exp\left\{\beta \int_{-T/2}^{T/2} r(t)s(t-mT)\mathsf{d}t\right\}.$$

A Physical Perspective ... (Cont'd)

Assume $s(t - mT)$ has duration Δ and divide $[-M, M]$ to $K = 2MT/\Delta$ subintervals $\mathcal{M}_i$.

ML estimation: find $\epsilon_i = \max_{m \in \mathcal{M}_i} \int_{-T/2}^{T/2} r(t)s(t - mT)\mathrm{d}t$, then $\max_i \epsilon_i$.

Define another paritition function

$$Z(\beta) = \sum_i e^{\beta \epsilon_i}.$$

The RV's $\{\epsilon_i\}$ are alternately independent, with a density known for some waveforms, e.g., rectangular pulses (Slepian '62, Shepp '66, Zakai & Ziv '69).

For T large, the tail is $\approx$ Gaussian.

The contribution $e^{\beta \epsilon_0}$ of the subinterval that includes the signal should be handled separately.

The Random Energy Model (REM)

The analysis of $Z(\beta)$ – very similar to that of the random energy model (REM) in statistical physics (Derrida ‘80,‘81):

$$Z(\beta) = \int \mathrm{d}\epsilon \cdot N(\epsilon) \cdot e^{\beta\epsilon},$$

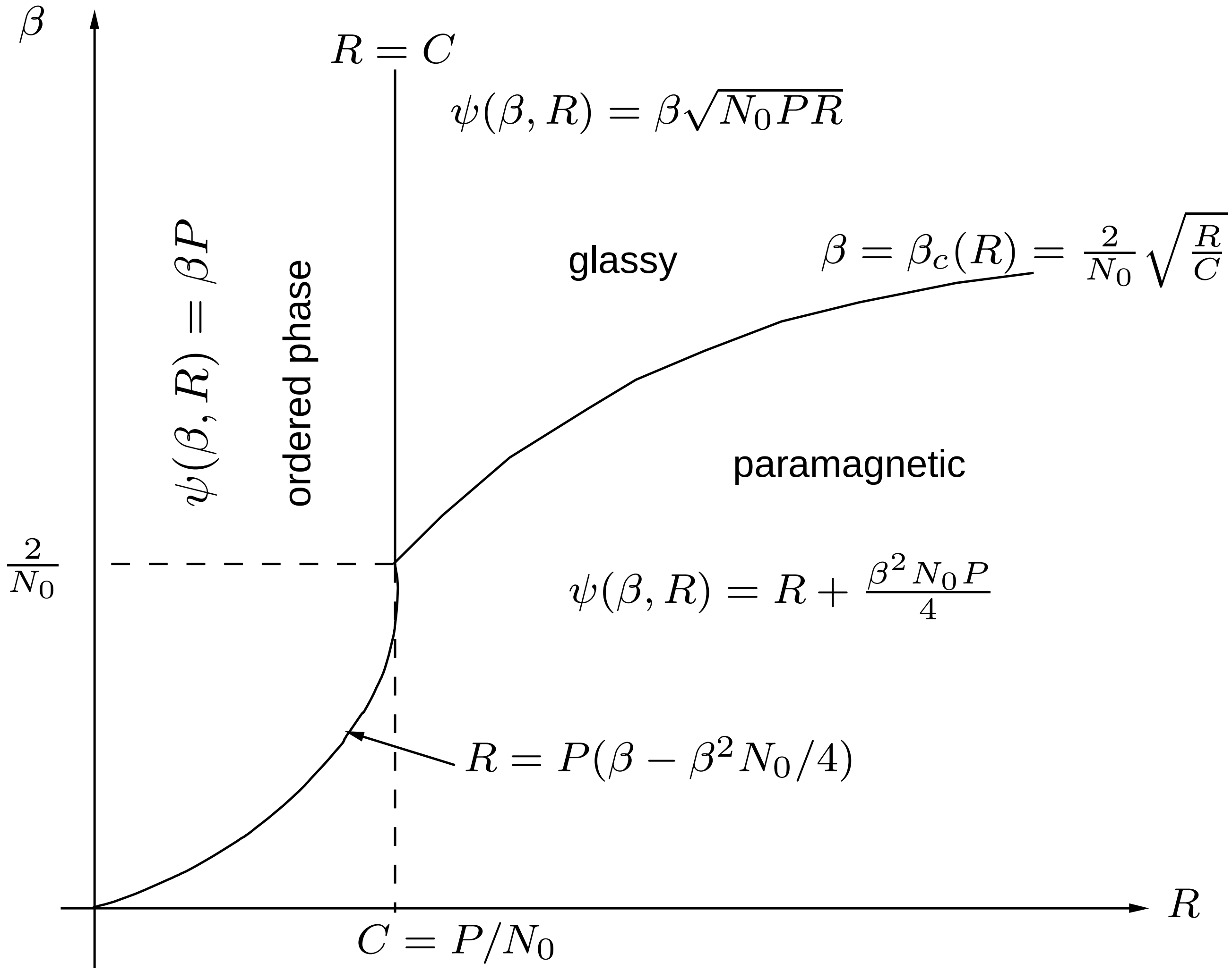
where typically

$$N(\epsilon)\mathrm{d}\epsilon \sim \begin{cases} 0 & f(\epsilon)\mathrm{d}\epsilon \ll 1/K \\ K \cdot f(\epsilon)\mathrm{d}\epsilon & \text{elsewhere} \end{cases}$$

For $W \sim e^{RT}$, we select $\Delta \sim e^{-RT} \implies K \sim e^{RT}$.

Accordingly, from now on, we denote the partition function by $Z(\beta, R)$ and define

$$\psi(\beta, R) = \lim_{T\to\infty} \frac{\ln Z(\beta, R)}{T}.$$


β
$R = C$
$\psi(\beta, R) = \beta\sqrt{N_0 P R}$
$\psi(\beta, R) = \beta P$
ordered phase
glassy
$\beta = \beta_c(R) = \frac{2}{N_0}\sqrt{\frac{R}{C}}$
paramagnetic
$\frac{2}{N_0}$
$\psi(\beta, R) = R + \frac{\beta^2 N_0 P}{4}$
$R = P(\beta - \beta^2 N_0/4)$
R
$C = P/N_0$

Some Extensions

- Mismatched estimation: suppose that $r(t)$ is correlated with $\tilde{s}(t-mT)$ instead of $s(t-mT)$: Same phase diagram, except that that C is degraded by a factor of ρ^2 and β by a factor of ρ, where ρ is the correlation between $s(\cdot)$ and $\tilde{s}(\cdot)$.
- Other pulse shapes: essentially the same results.

Joint ML Estimation of Amplitude and Delay

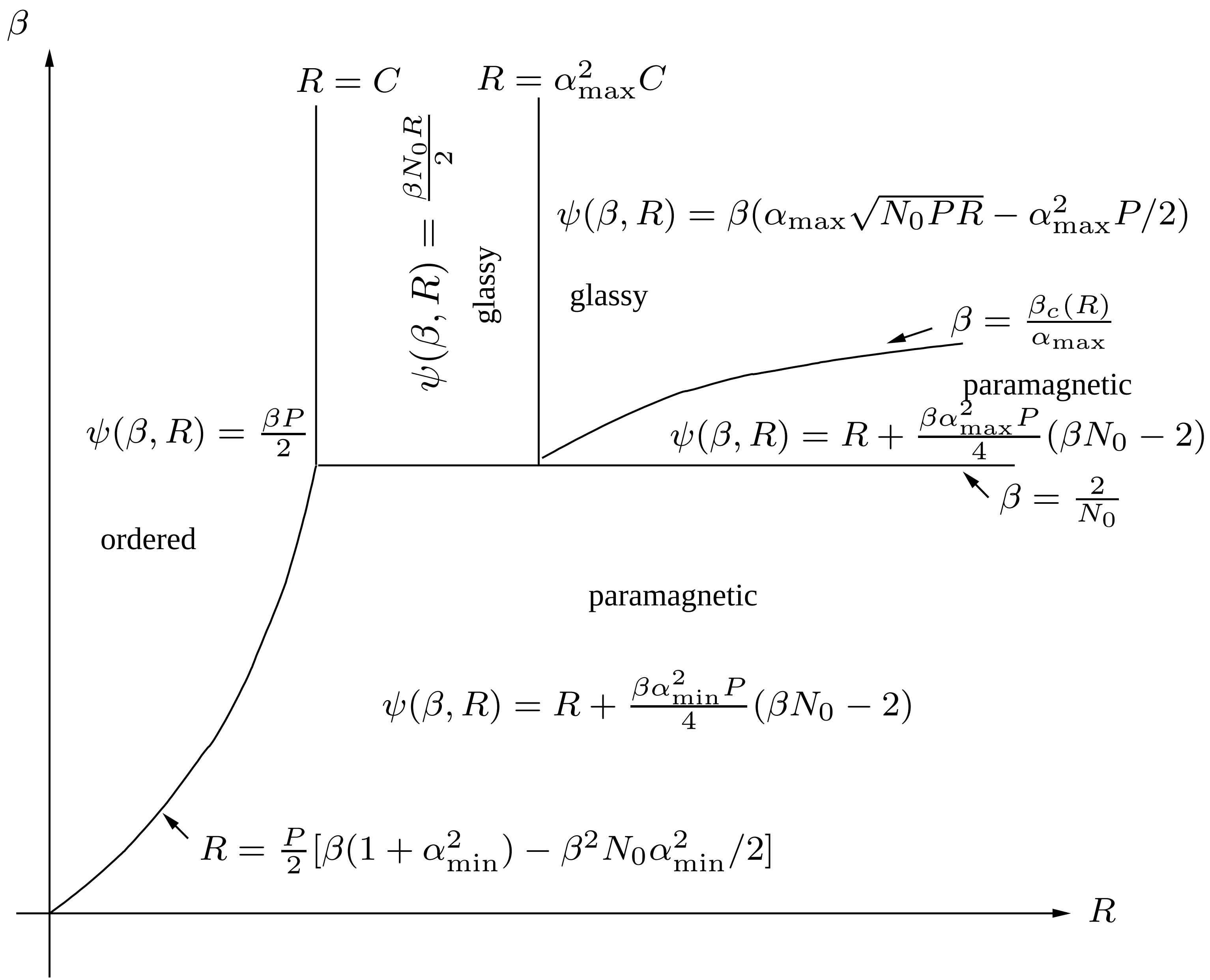
Consider now the model

$$r(t) = \alpha \cdot s(t - mT) + n(t),$$

where now both m and α have to be estimated. It is assumed that $\alpha \in [\alpha_{\min}, \alpha_{\max}]$.

- The parameter α alone does not contribute any phase transitions.
- The parameter m alone generates three phases.

Q: How many phases would there be in the joint estimation of m and α?



β
$R = C$
$R = \alpha_{\max}^2 C$
$\psi(\beta, R) = \frac{\beta N_0 R}{2}$
glassy
$\psi(\beta, R) = \beta(\alpha_{\max}\sqrt{N_0 P R} - \alpha_{\max}^2 P/2)$
glassy
$\beta = \frac{\beta_c(R)}{\alpha_{\max}}$
paramagnetic
$\psi(\beta, R) = \frac{\beta P}{2}$
$\psi(\beta, R) = R + \frac{\beta \alpha_{\max}^2 P}{4}(\beta N_0 - 2)$
$\beta = \frac{2}{N_0}$
ordered
paramagnetic
$\psi(\beta, R) = R + \frac{\beta \alpha_{\min}^2 P}{4}(\beta N_0 - 2)$
$R = \frac{P}{2}[\beta(1 + \alpha_{\min}^2) - \beta^2 N_0 \alpha_{\min}^2/2]$
R

Discussion and Conclusion

- The behavior is much more complicated than when only m should be estimated.
- One ordered phase (non–anomalous errors) and four anomalous phases.
- Although α alone does not generate phase transitions, its interaction with m generates more phases than those of m alone.
- Anomalies in α have a different behavior.

The physical point of view helps to gain insight on the behavior.