Fast evaluation and root finding for polynomials with floating-point coefficients

Guillaume Moroz, Rémi Imbach

9 mars 2023

Floating-point arithmetic

Representation

Light-year: $9\,460\,730\,472\,580\,800$ m

$$\underbrace{9.460}_{\textbf{mantissa}} \cdot 10^{\overbrace{15}^{\textbf{exponent}}} \text{ m}$$

number of digits $< m$ absolute value $< \tau$

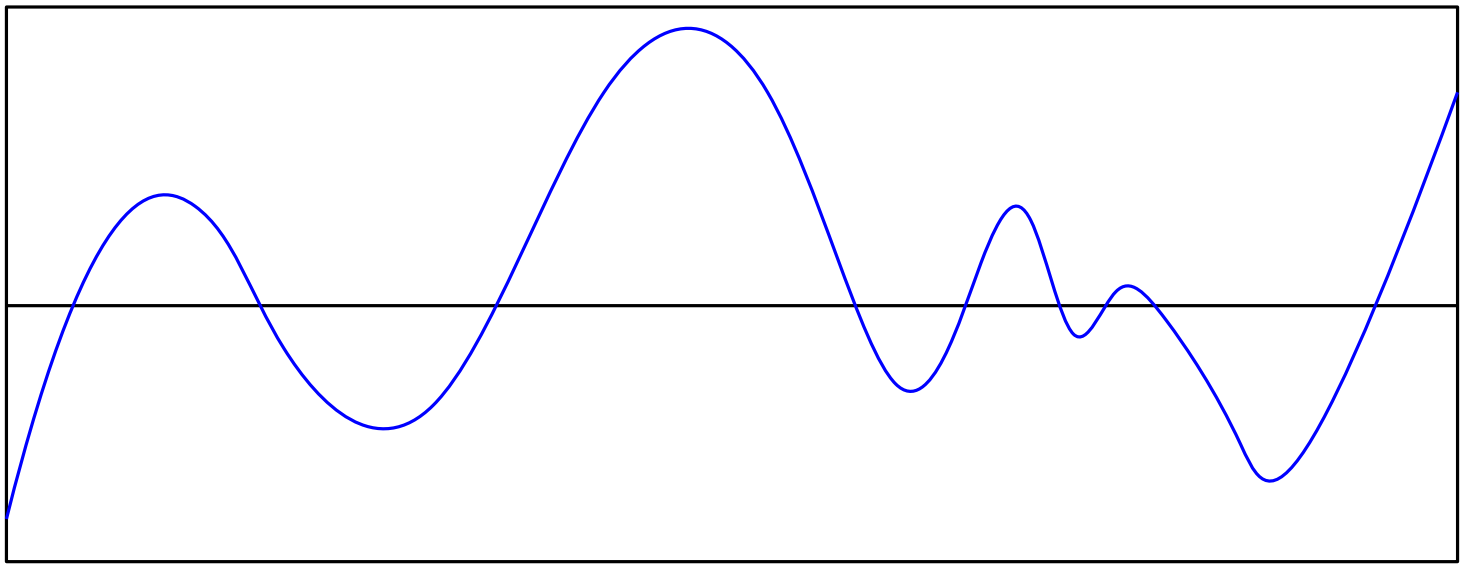
Polynomial evaluation

$$f(z) = f_0 + \cdots + f_d z^d$$

- Bit complexity : $\widetilde{O}(d(m + \log \tau))$
- Error: $O(d2^{-m})\widetilde{f}(|z|)$ where $\widetilde{f}(|z|) = \sum |f_j||z|^j$

Main result

$$f(z) = f_0 + \cdots + f_d z^d$$



Theorem (Piecewise polynomial approximation)

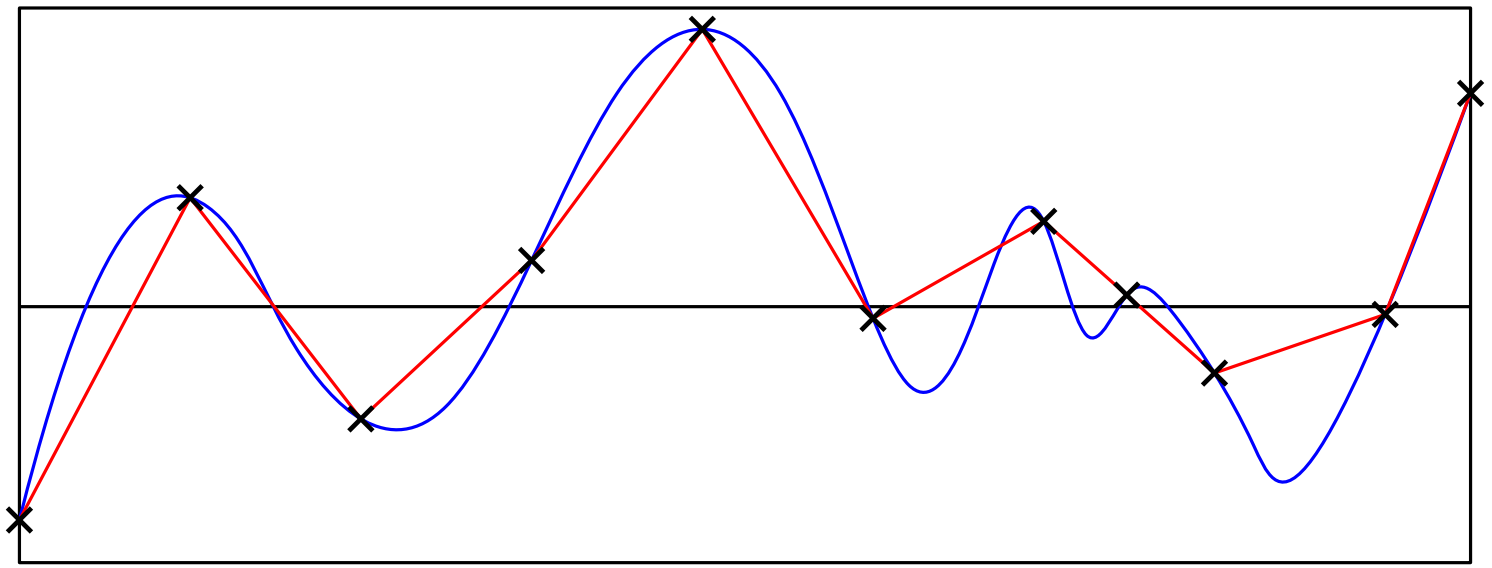
It is possible to compute all g_n of degree m satisfying

$$|f(z) - g_n(z)| < 2^{-m} \widetilde{f}(|z|)$$

in $\widetilde{O}(d(m + \log \tau))$ bit operations

Main result

$$f(z) = f_0 + \cdots + f_d z^d$$



Theorem (Piecewise polynomial approximation)

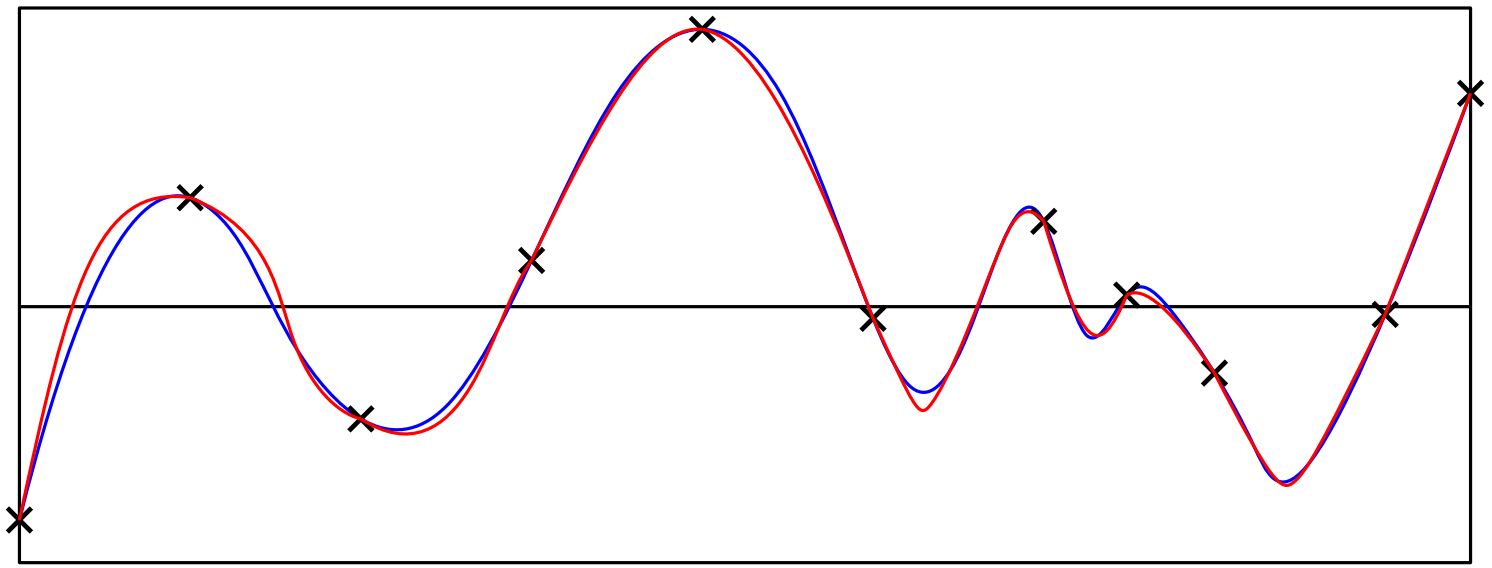
It is possible to compute all g_n of degree m satisfying

$$|f(z) - g_n(z)| < 2^{-m} \widetilde{f}(|z|)$$

in $\widetilde{O}(d(m + \log \tau))$ bit operations

Main result

$$f(z) = f_0 + \cdots + f_d z^d$$



Theorem (Piecewise polynomial approximation)

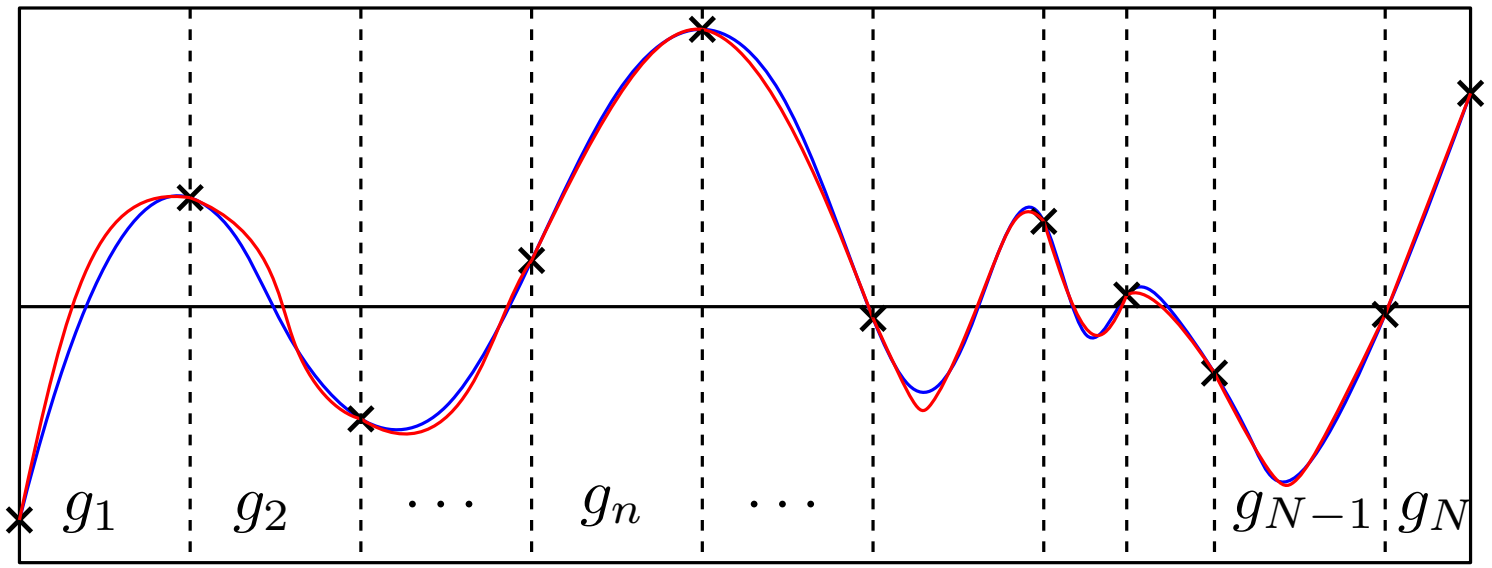
It is possible to compute all g_n of degree m satisfying

$$|f(z) - g_n(z)| < 2^{-m} \widetilde{f}(|z|)$$

in $\widetilde{O}(d(m + \log \tau))$ bit operations

Main result

$$f(z) = f_0 + \cdots + f_d z^d$$



Theorem (Piecewise polynomial approximation)

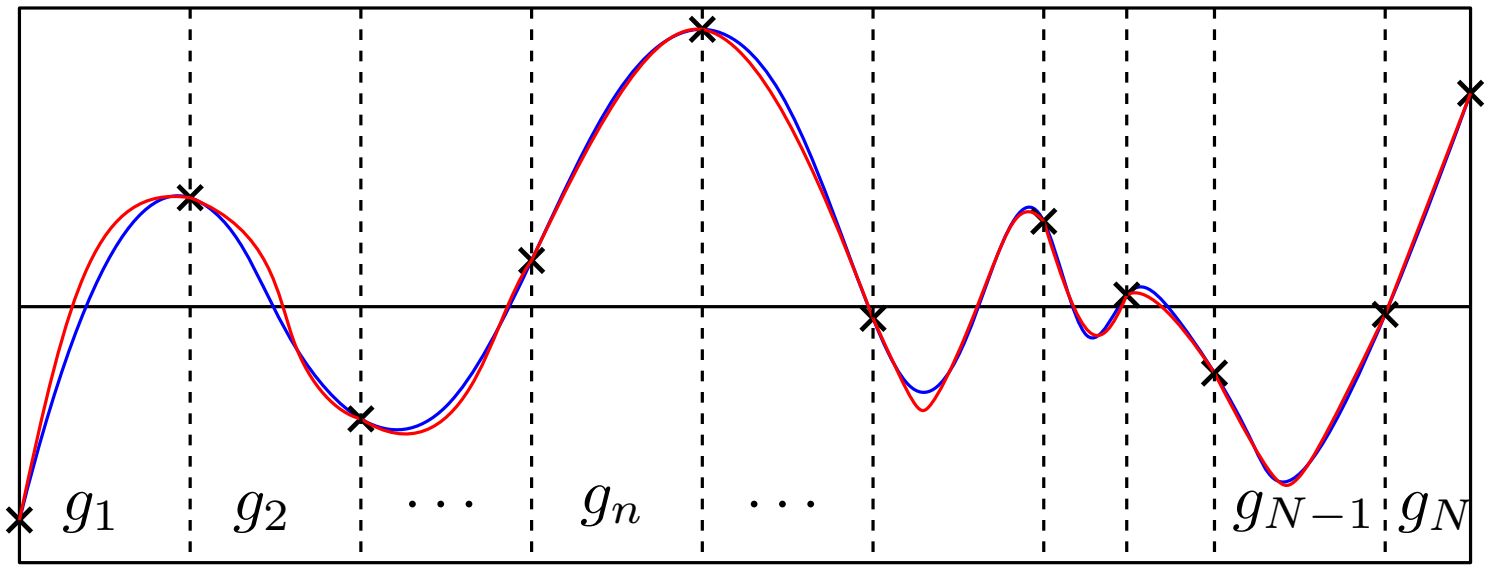
It is possible to compute all g_n of degree m satisfying

$$|f(z) - g_n(z)| < 2^{-m} \widetilde{f}(|z|)$$

in $\widetilde{O}(d(m + \log \tau))$ bit operations

Corollary

$$f(z) = f_0 + \cdots + f_d z^d \qquad m > \log d$$



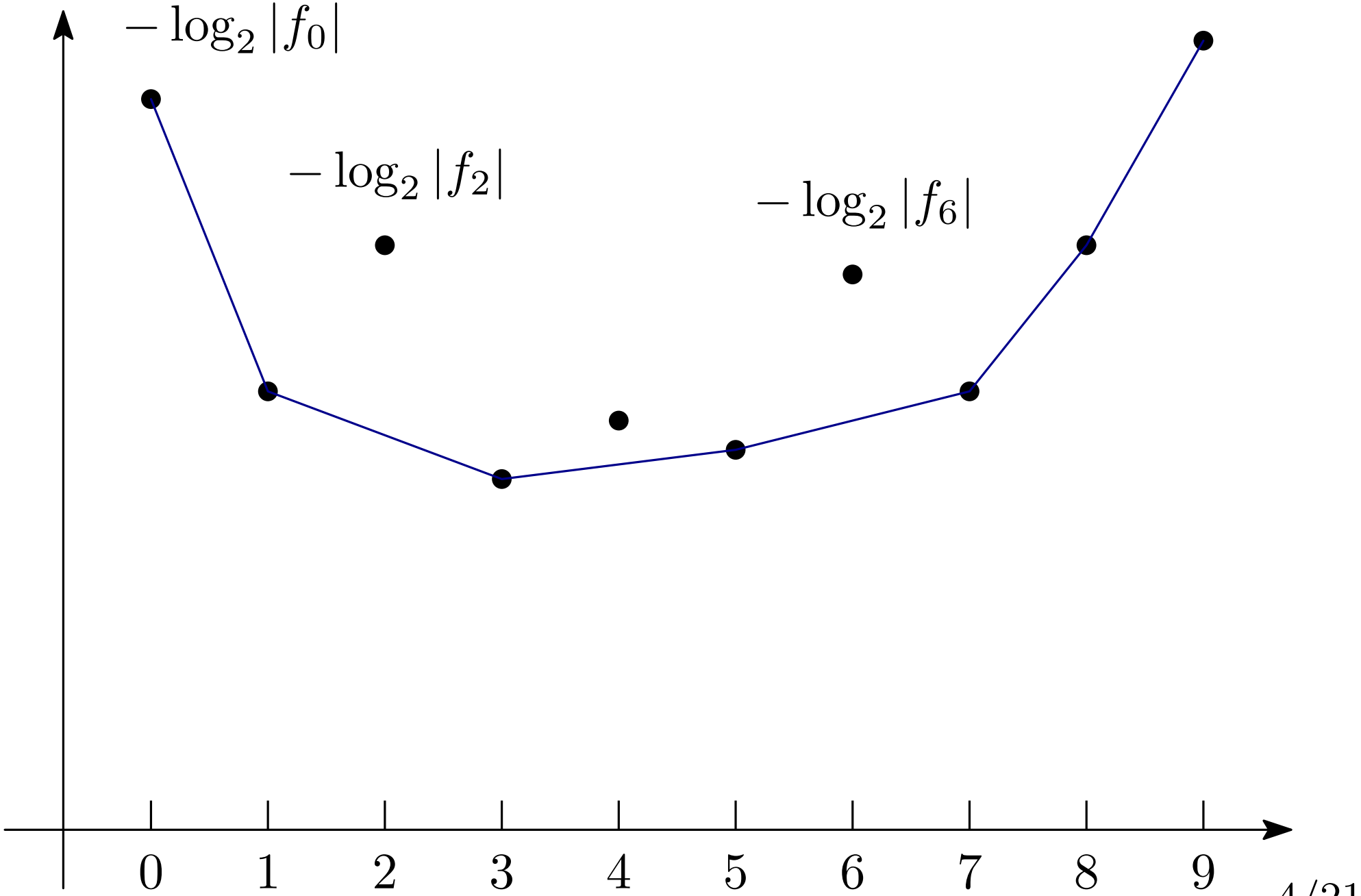
Fast evaluation

With error $2^{-m}\widetilde{f}(|z|)$
in
$\widetilde{O}(m(m + \log \tau))$

Fast root finding

All roots with m bits
in
$\widetilde{O}(d(m + \log \tau + \log \kappa))$

Newton polygon



Newton polygon

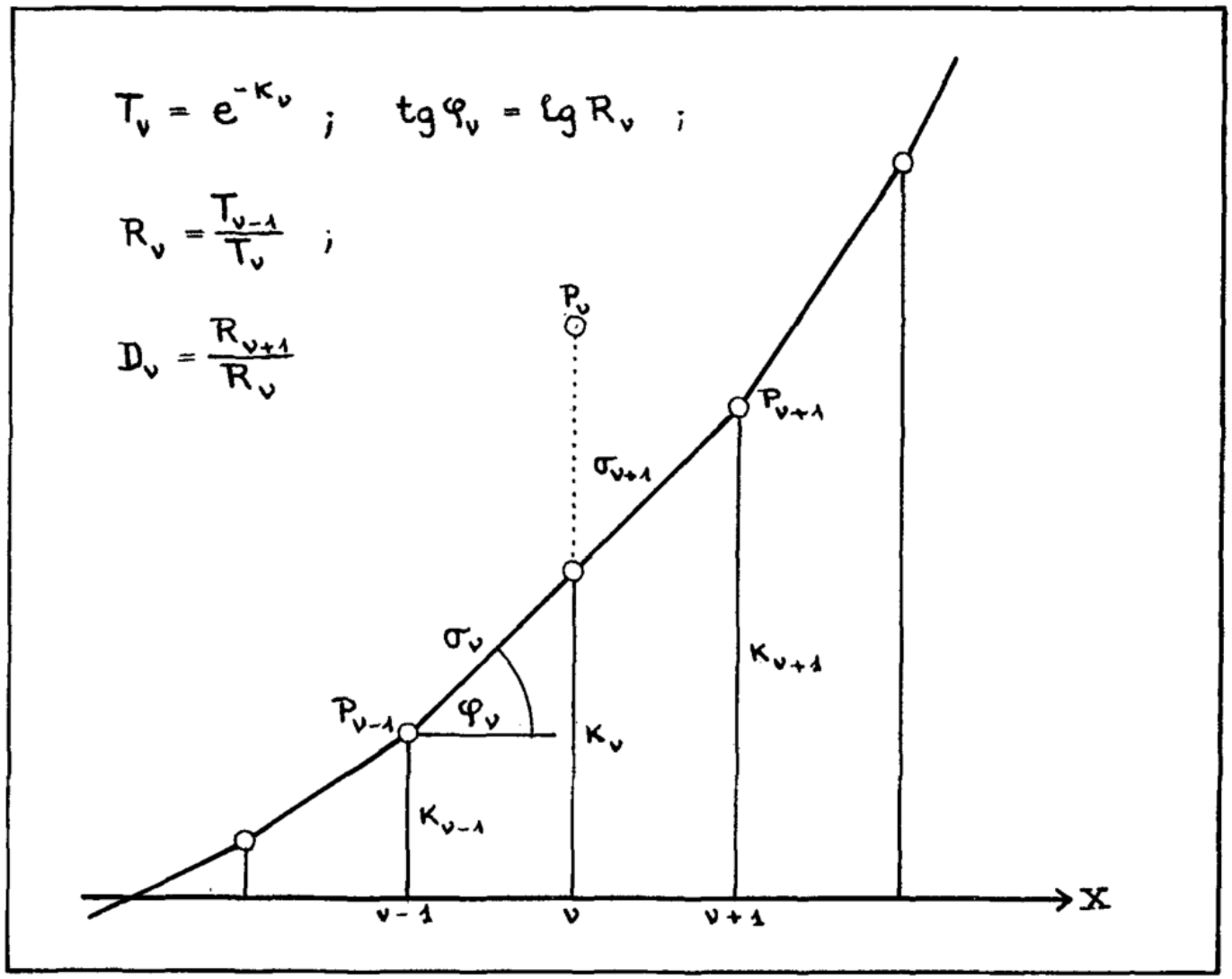
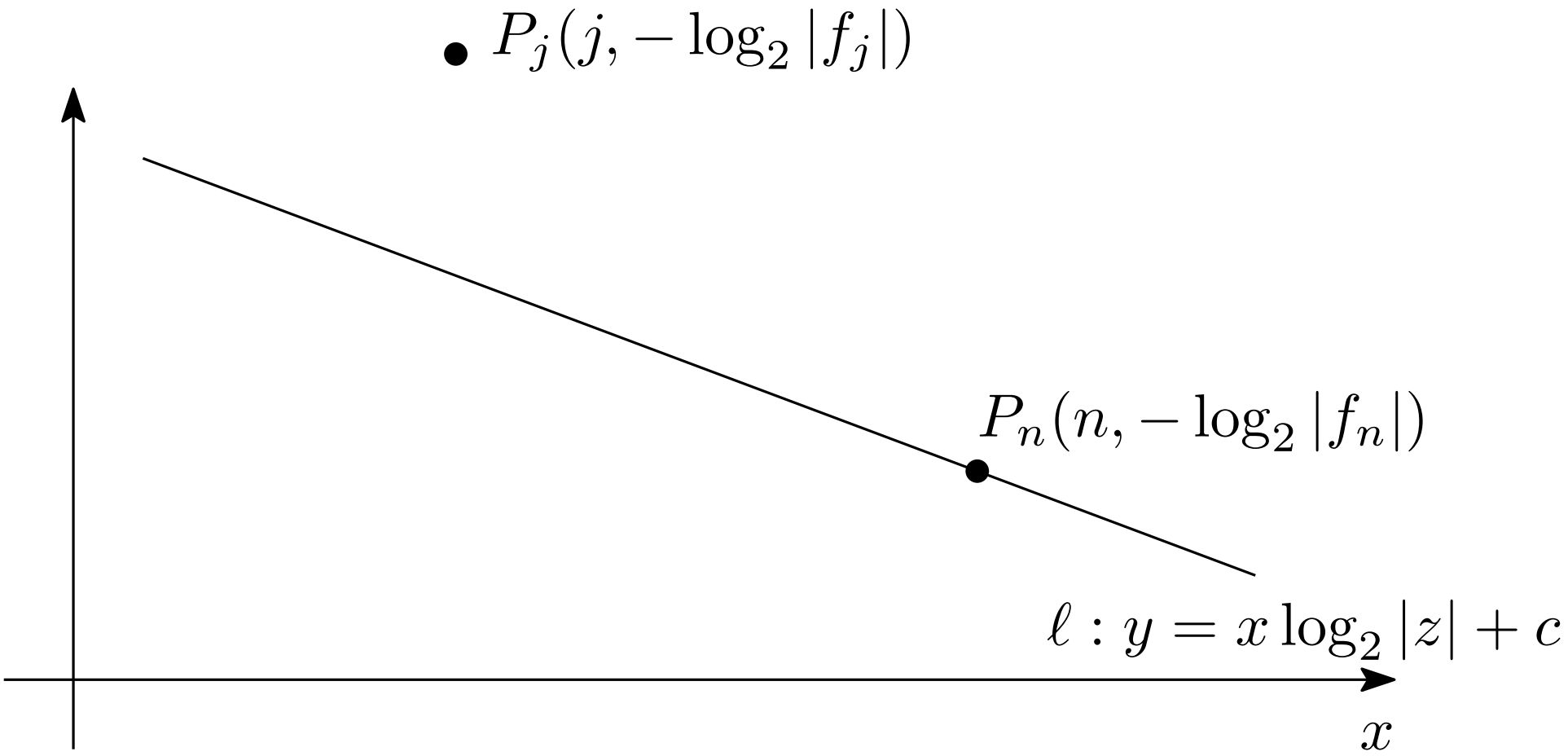


Fig. 1.

Alexandre Ostrowski. *Recherches sur la méthode de graeffe et les zéros des polynomes et des séries de laurent.* Acta Mathematica, 1940

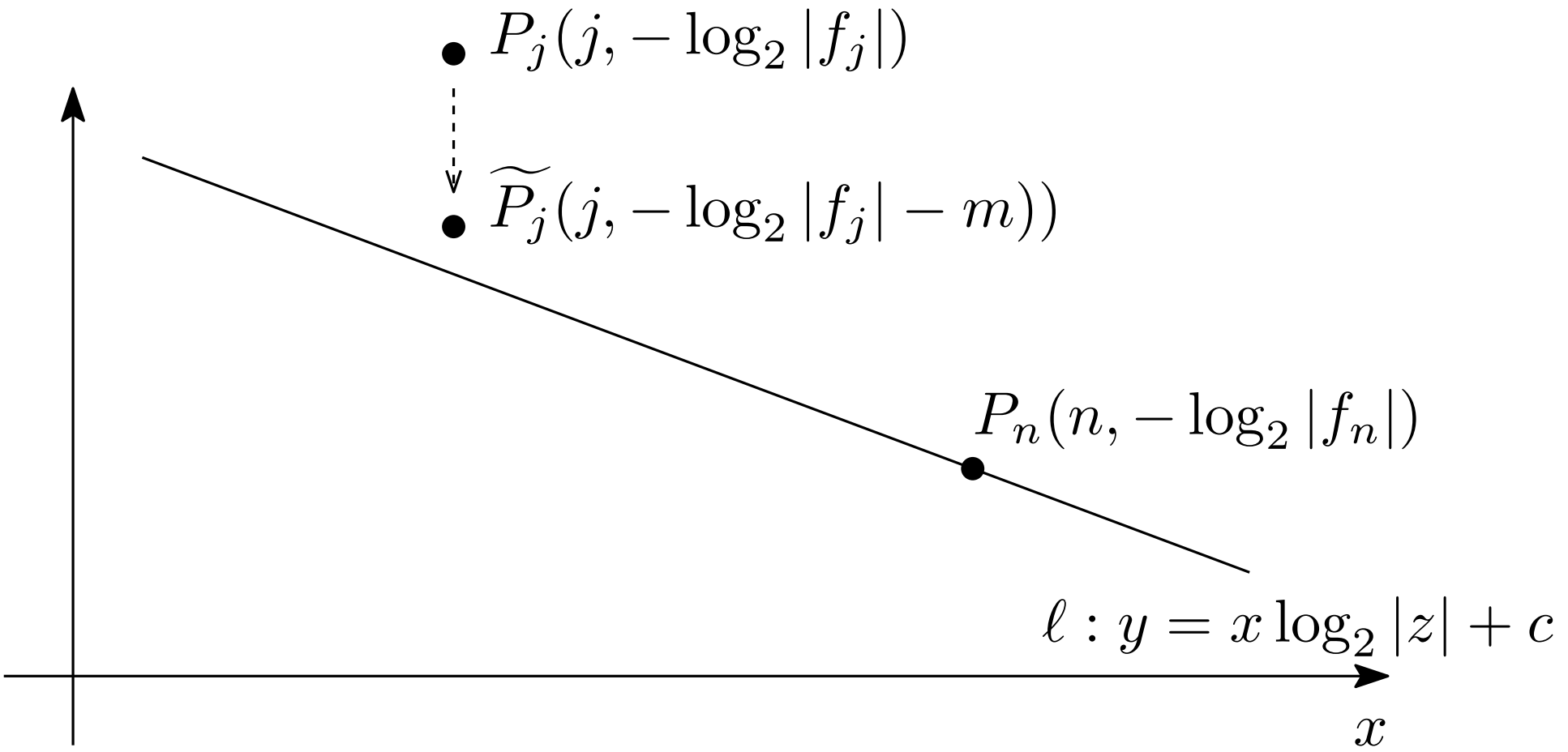
Newton polygon



Dominant monomial

P_j above ℓ $\iff$ $|f_j||z|^j <$ $|f_n||z|^n$

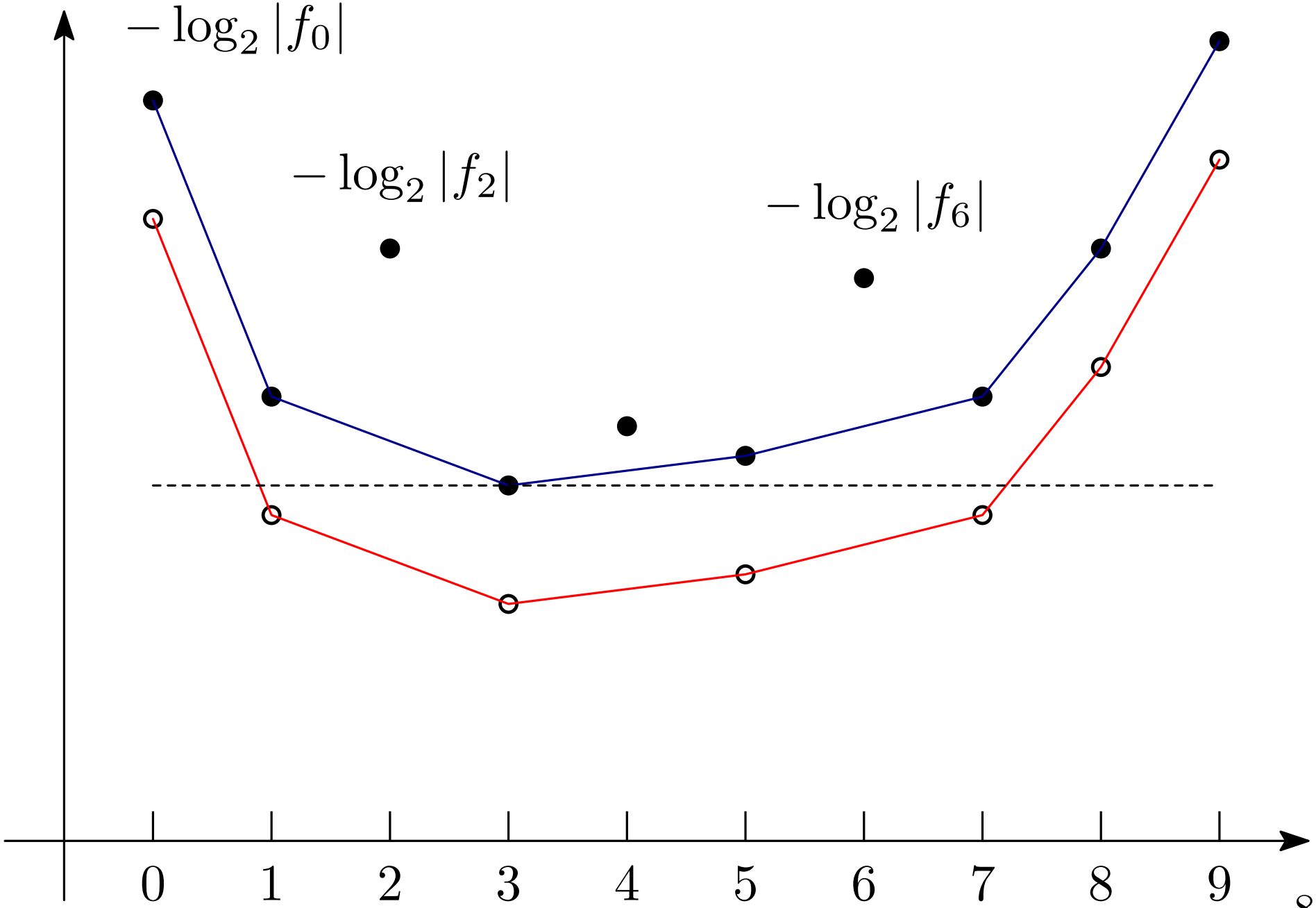
Newton polygon



Dominant monomial

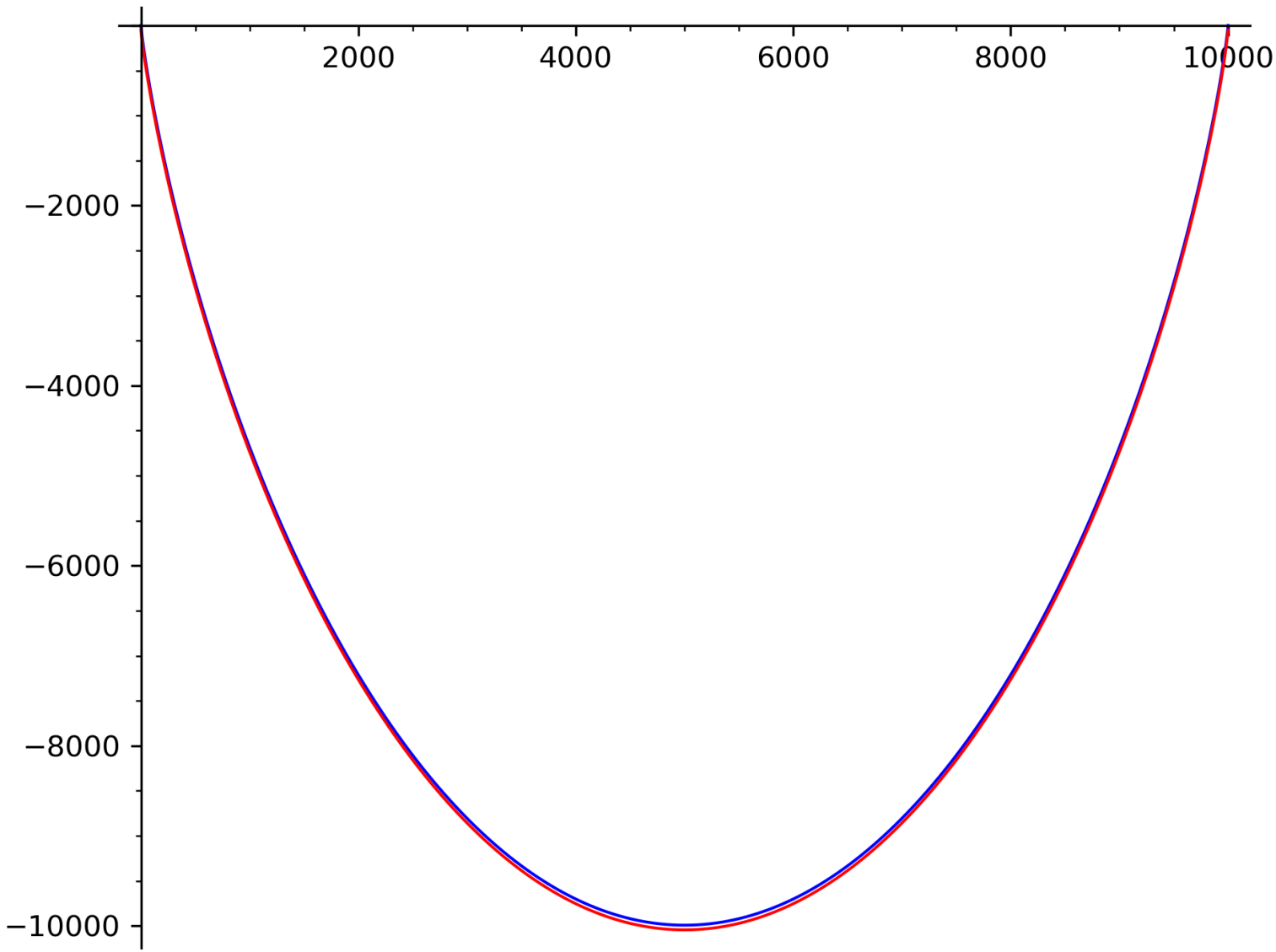
P_j above ℓ	$\iff$	$\lvert f_j\rvert\lvert z\rvert^j < \lvert f_n\rvert\lvert z\rvert^n$
$\widetilde{P_j}$ above ℓ	$\iff$	$\lvert f_j\rvert\lvert z\rvert^j < 2^{-m}\lvert f_n\rvert\lvert z\rvert^n$

Newton polygon



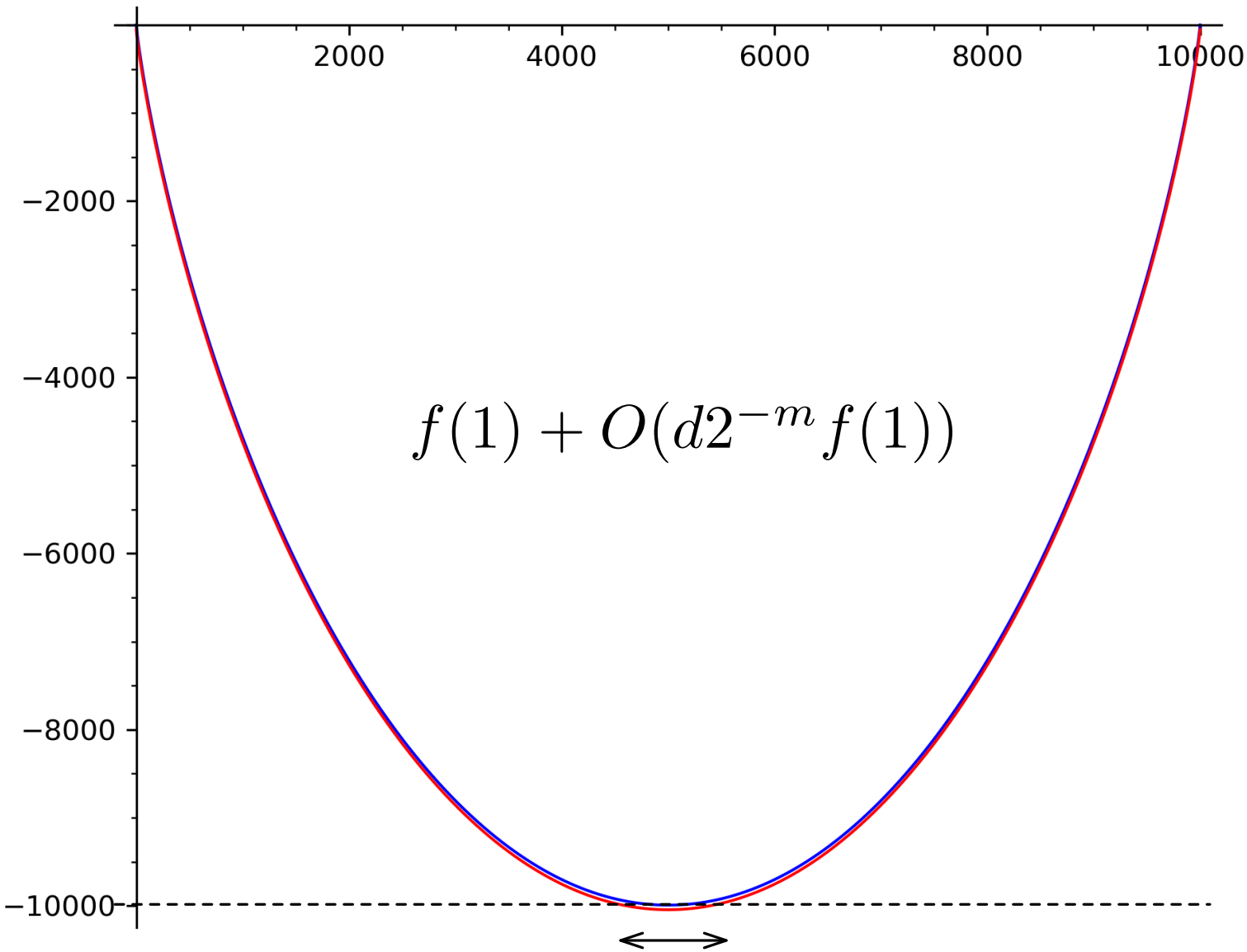
Example: selecting monomials

$$f(z) = (1+z)^{10000} = \sum \binom{10000}{j} z^j$$



Example: selecting monomials

$$f(z) = (1+z)^{10000} = \sum \binom{10000}{j} z^j$$



Example: Taylor expansion

$$1$$

$$10\,000z$$

$$\vdots$$

$$\binom{10\,000}{\ell} z^{\ell}$$

$$\binom{10\,000}{j} z^{j}$$

$$\binom{10\,000}{u} z^{u}$$

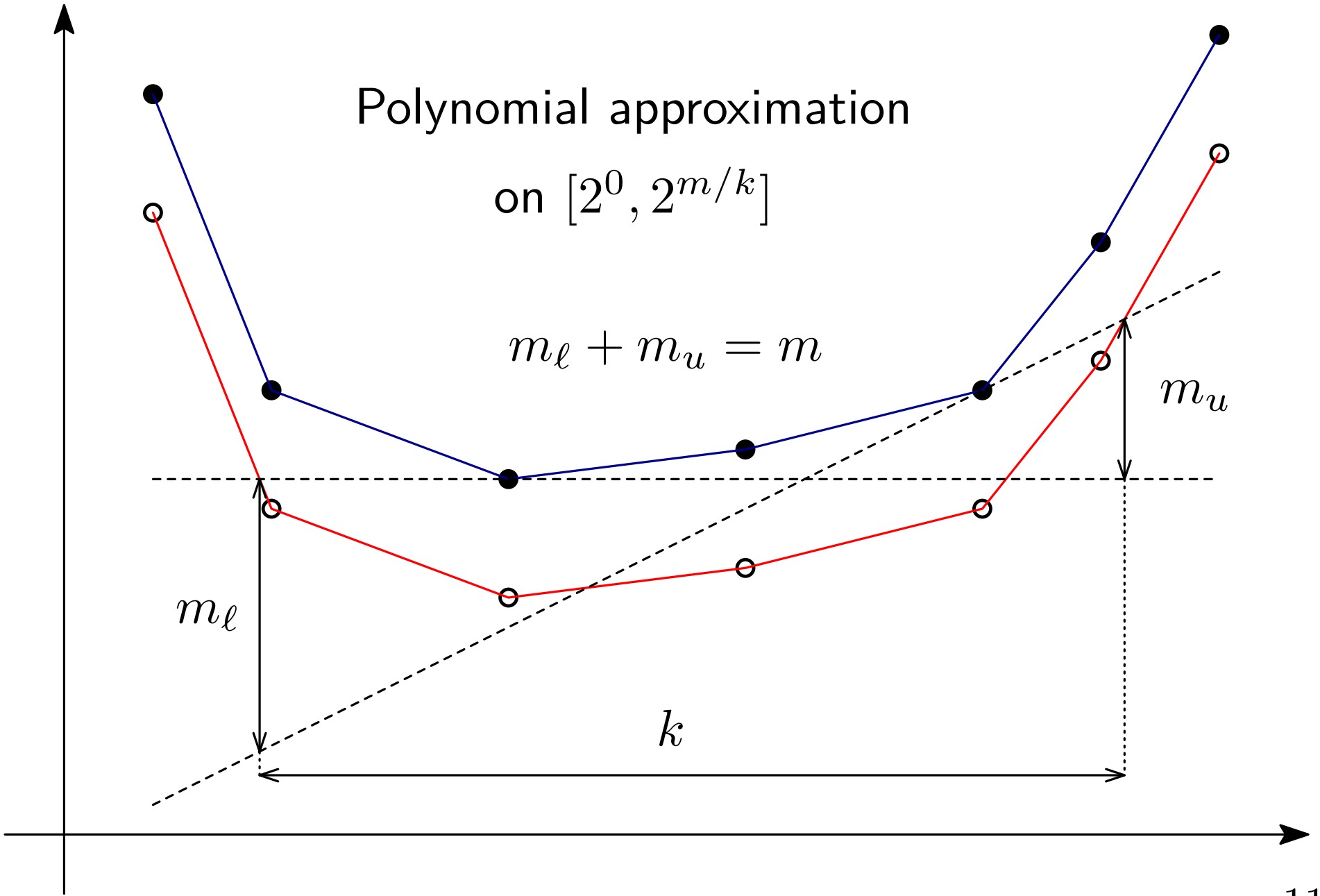
$$\vdots$$

$$z^{10\,000}$$

Example: Taylor expansion

$$
\begin{array}{l}
1 \qquad\qquad (1+\varepsilon t)^k = 1 + \cdots + \binom{k}{m}\varepsilon^m t^m + \mathcal{O}\left(\frac{k}{m}\varepsilon\right)^m \\
\\
10\,000z \\
\\
\quad \vdots \\
\\
k \left\{ \begin{array}{ll}
\binom{10\,000}{\ell} z^\ell & = \binom{10\,000}{\ell} z^\ell \\
\binom{10\,000}{j} z^j & = \binom{10\,000}{j} z^\ell (1+\varepsilon t)^{j-\ell} \\
\binom{10\,000}{u} z^u & = \binom{10\,000}{u} z^\ell (1+\varepsilon t)^{u-\ell}
\end{array} \right. \\
\\
\quad \vdots \\
\\
z^{10\,000}
\end{array}
$$

Algorithm: one interval



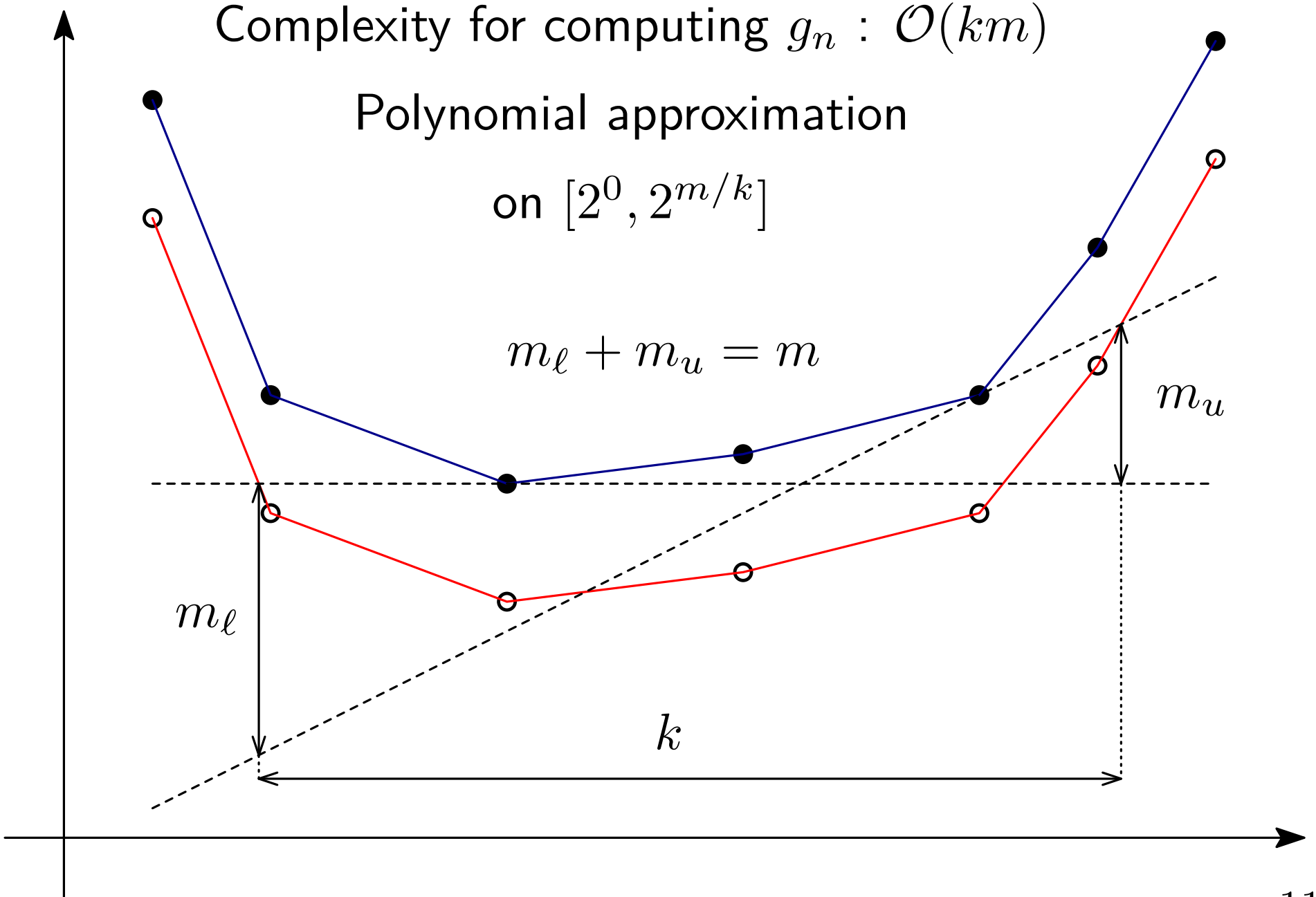
Algorithm: one interval

Complexity for computing g_n : $\widetilde{\mathcal{O}}(km)$

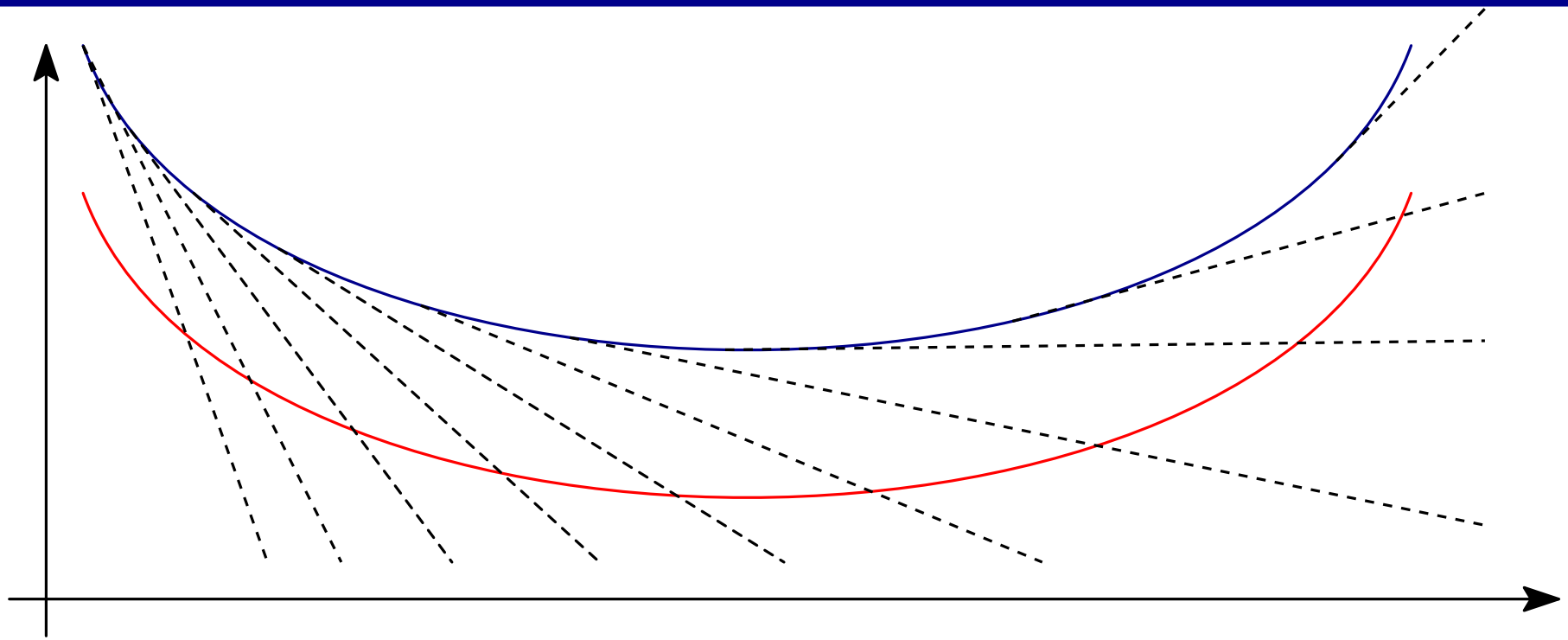
Polynomial approximation

on $[2^0, 2^{m/k}]$

$$m_\ell + m_u = m$$



Algorithm: all intervals



Approximation algorithm

$\textsc{While}\ a_j < a_{max}$

$a_{j+1} = a_j + \frac{m}{k_j}$

$g_{j+1} =$ Taylor approximation of degree $\mathcal{O}(m)$ on the interval $[2^{a_j}, 2^{a_{j+1}}]$

Complexity

Theorem

The piecewise approximation algorithm costs

$$\widetilde{\mathcal{O}}\left(\sum k_j m\right) \in \widetilde{\mathcal{O}}(dm) \text{ bit-operations}$$

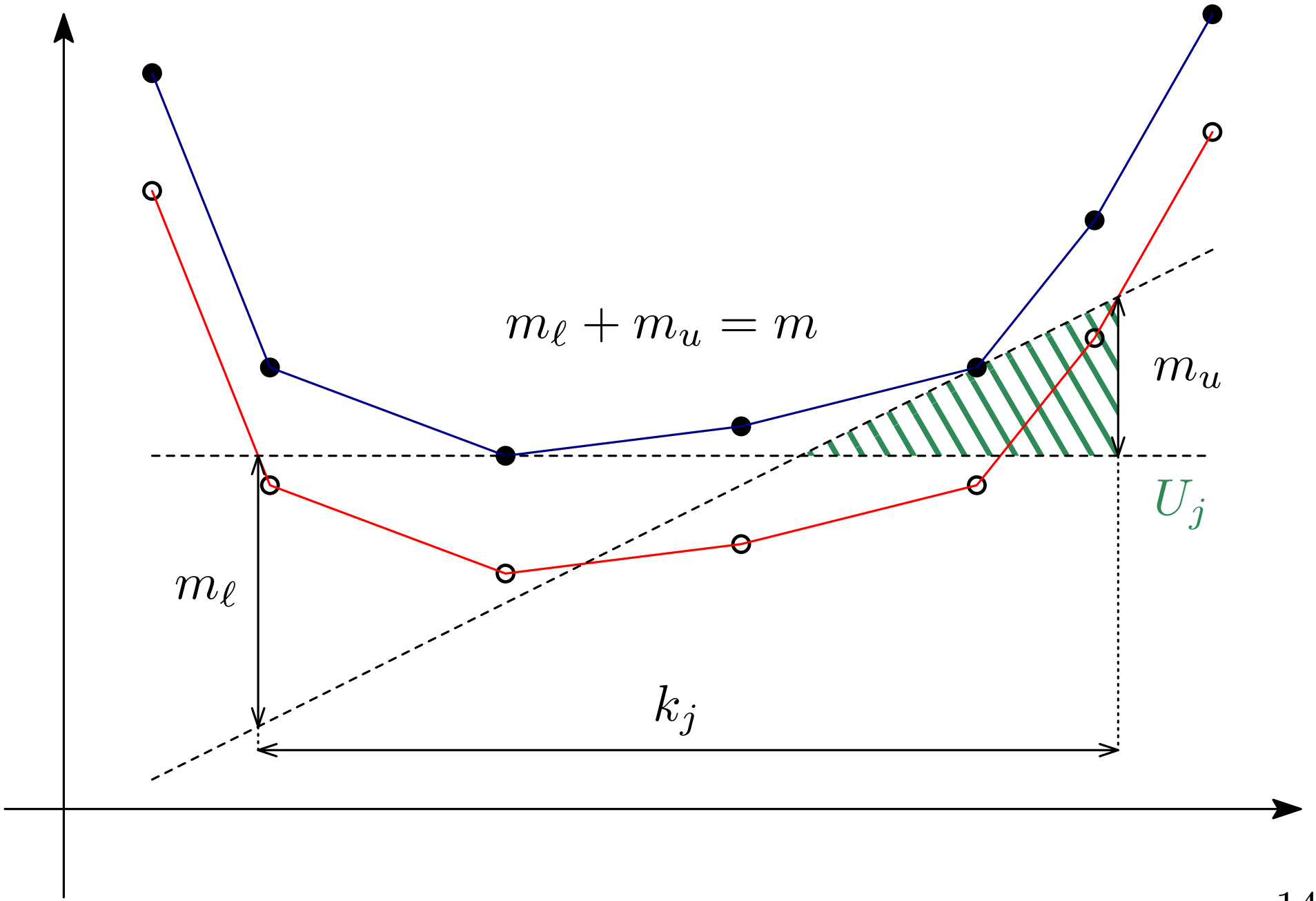
Approximation algorithm

$\textsc{While}\ a_j < a_{max}$

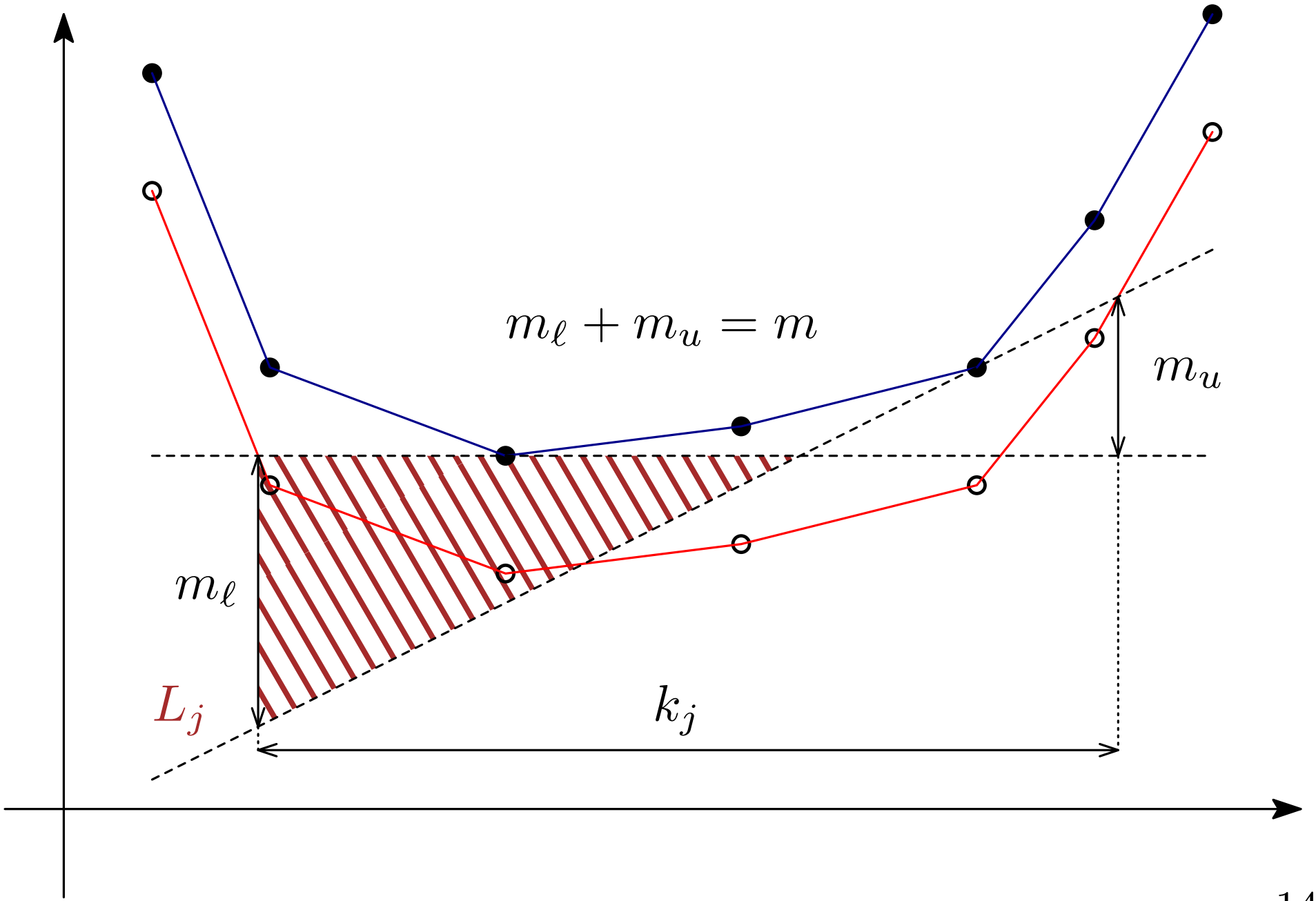
$a_{j+1} = a_j + \frac{m}{k_j}$

$g_{j+1} =$ Taylor approximation of degree $\mathcal{O}(m)$ on the interval $\left[2^{a_j}, 2^{a_{j+1}}\right]$

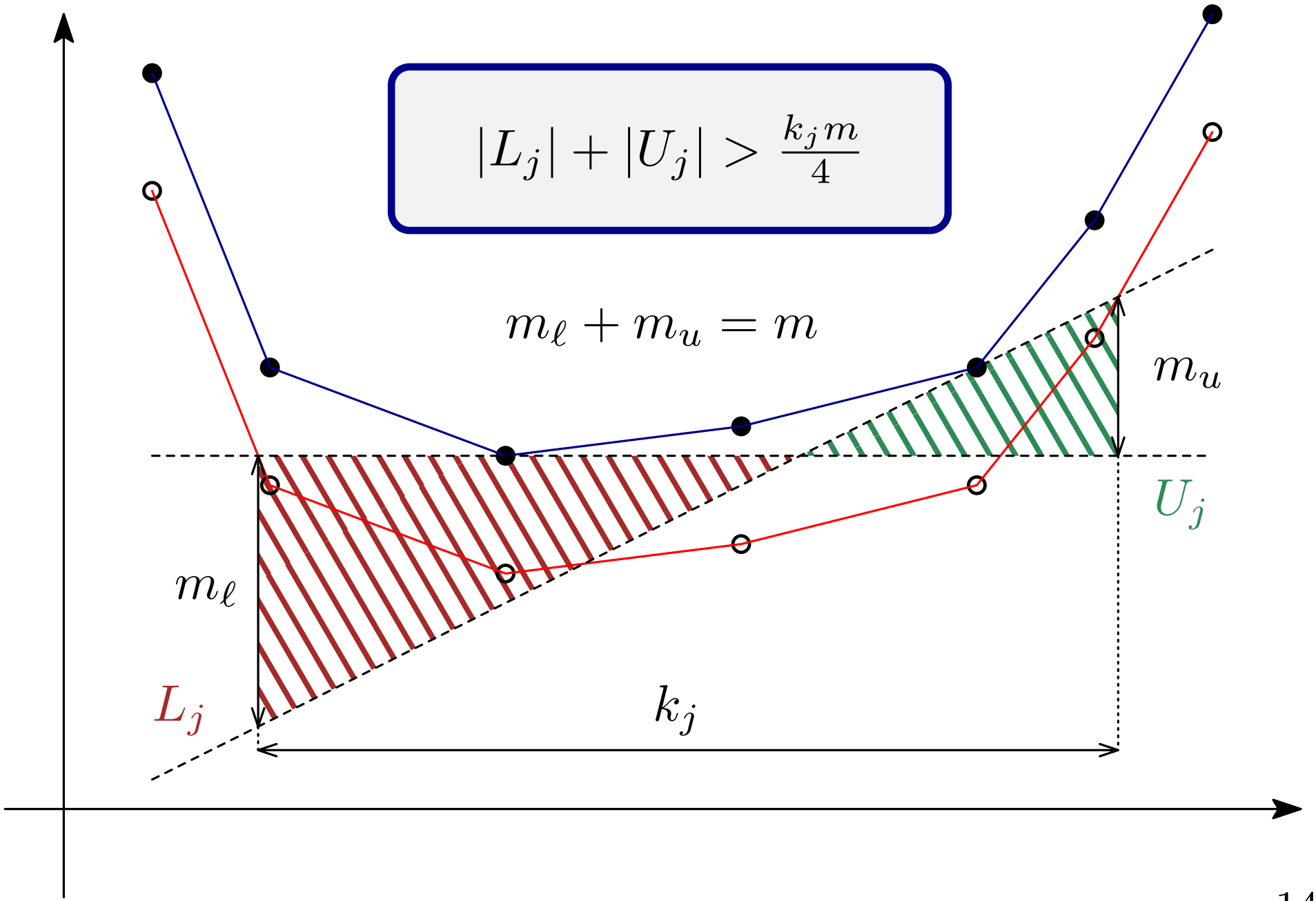
Proof



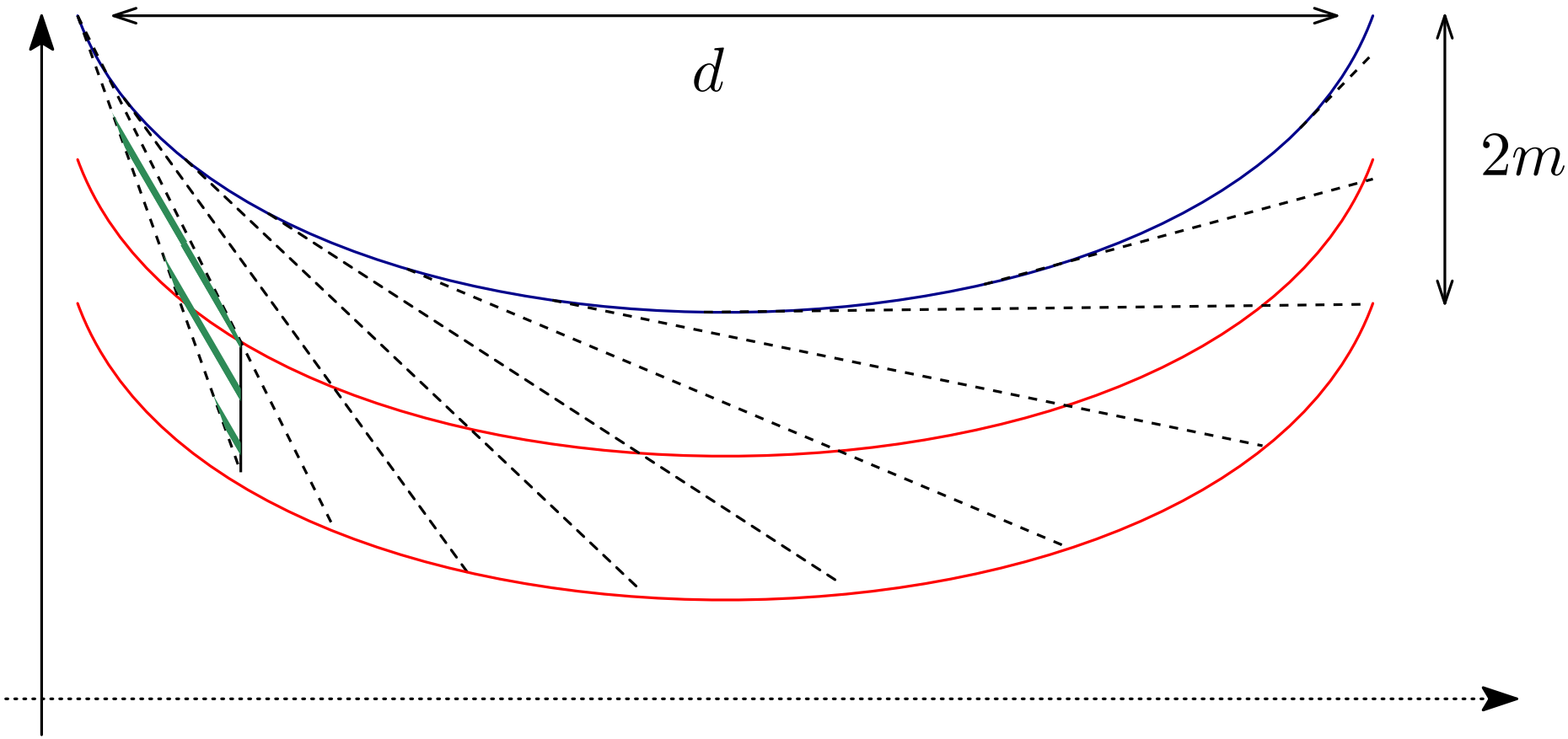
Proof



Proof



Proof

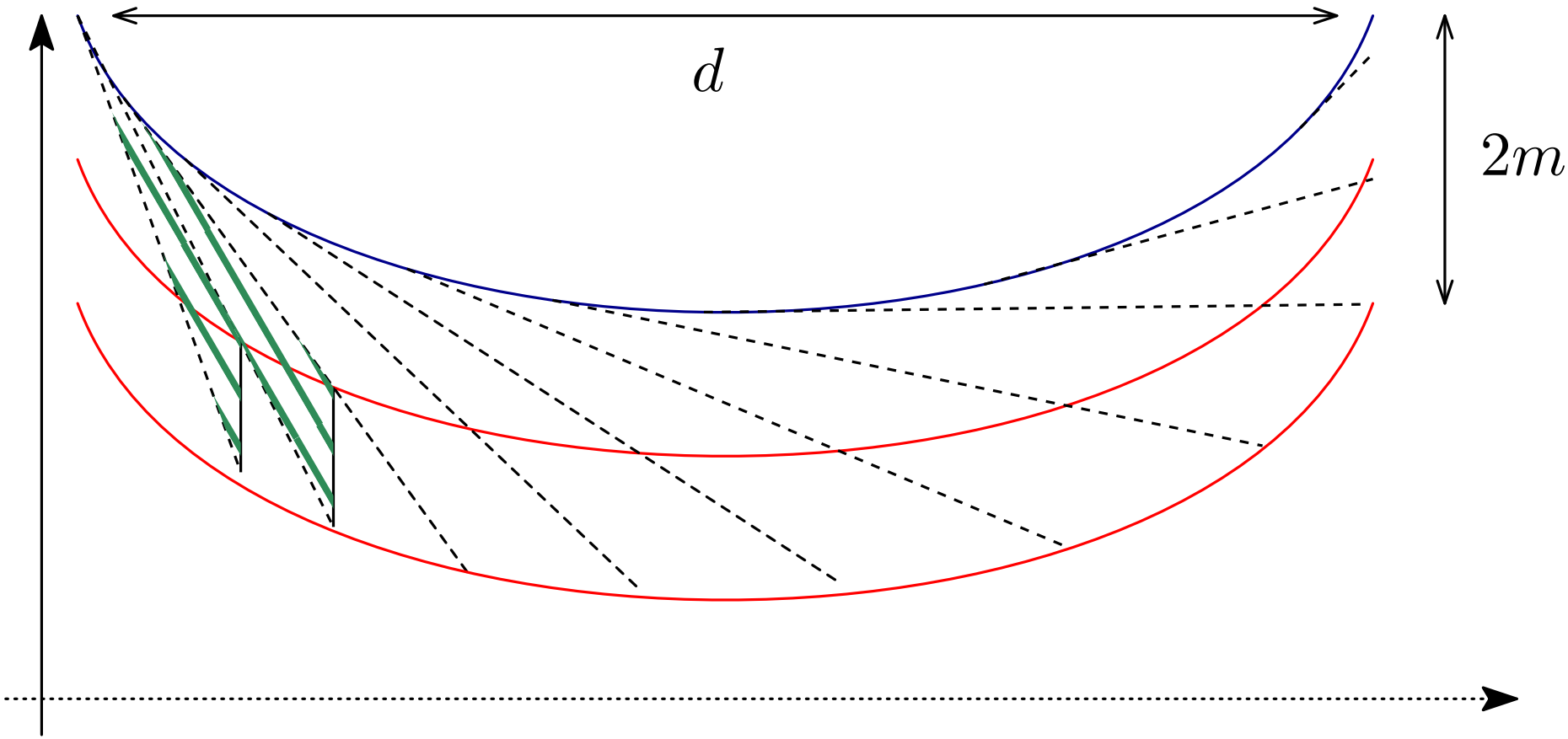


U_j piecewise disjoint

Union of U_j in B_{2m}

$$\sum |U_j| < 2dm$$

Proof

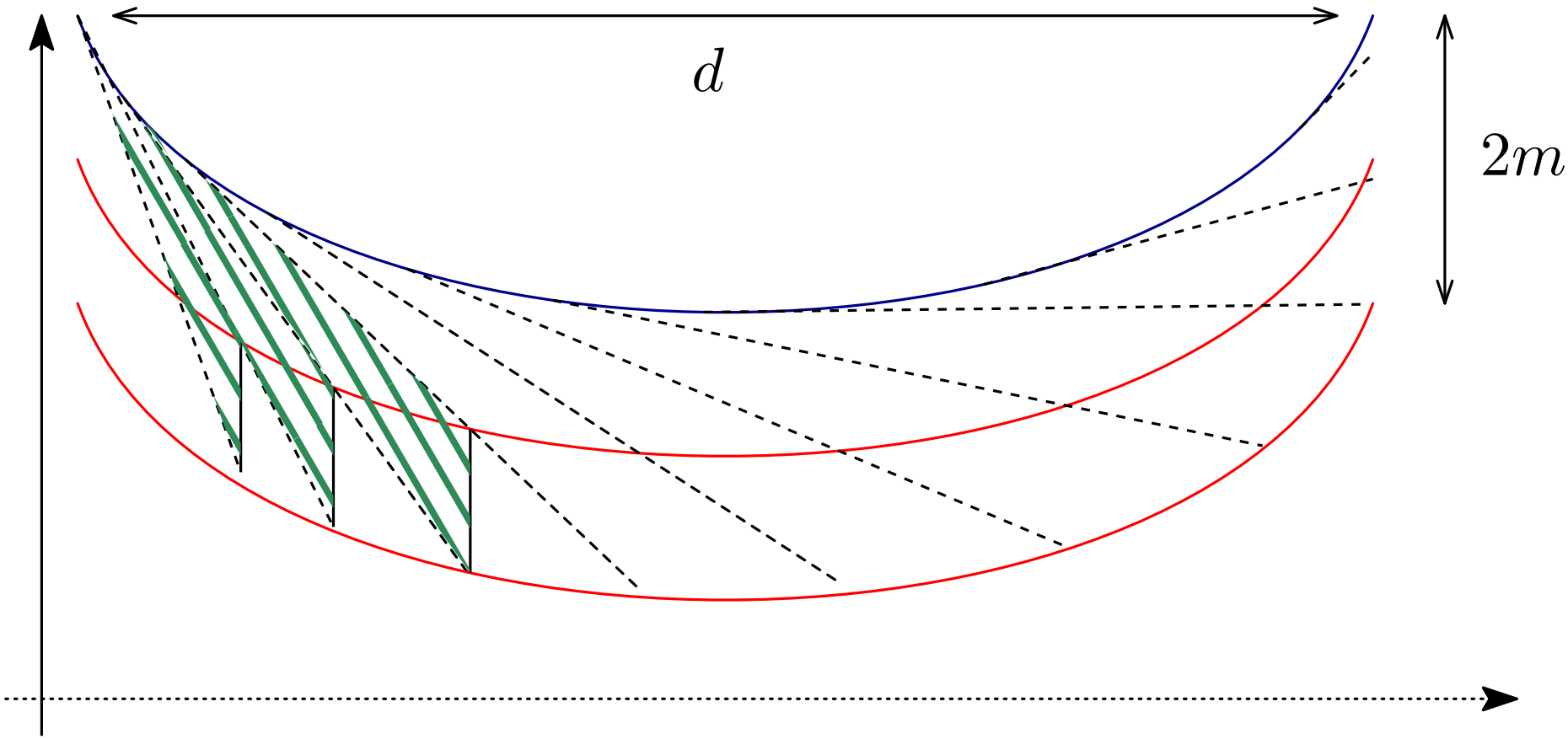


U_j piecewise disjoint

Union of U_j in B_{2m}

$$\sum |U_j| < 2dm$$

Proof

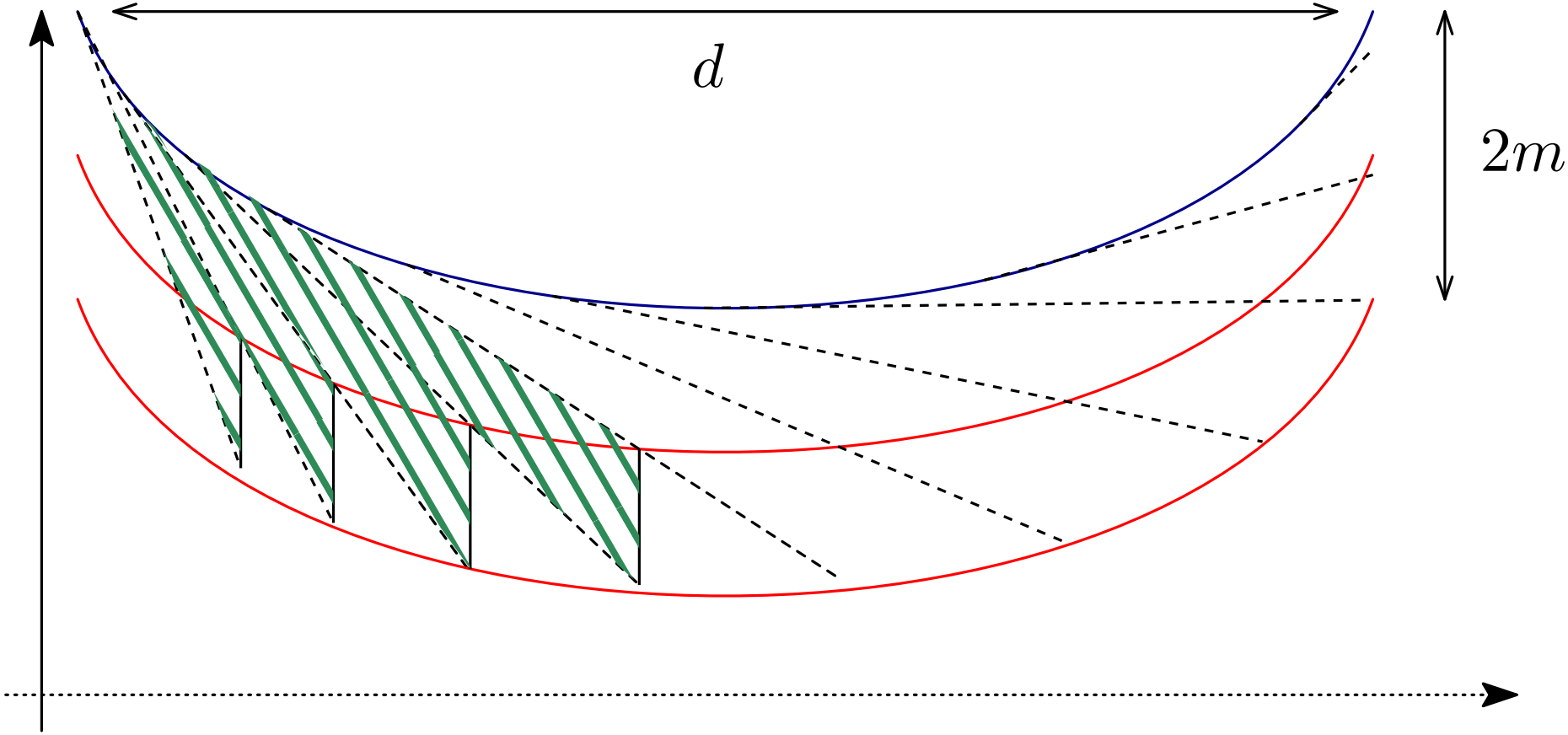


U_j piecewise disjoint

Union of U_j in B_{2m}

$$\sum |U_j| < 2dm$$

Proof

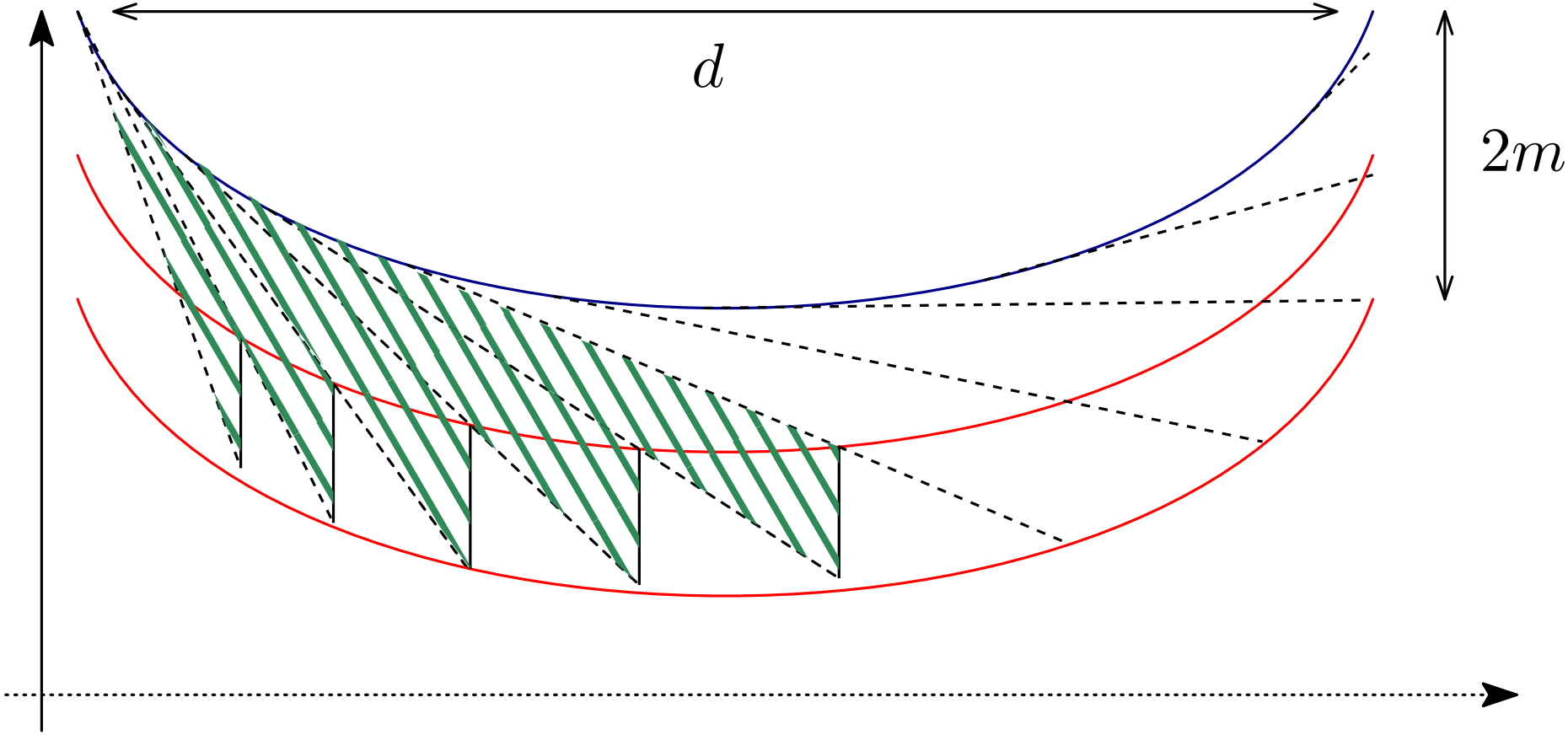


U_j piecewise disjoint

Union of U_j in B_{2m}

$$\sum |U_j| < 2dm$$

Proof

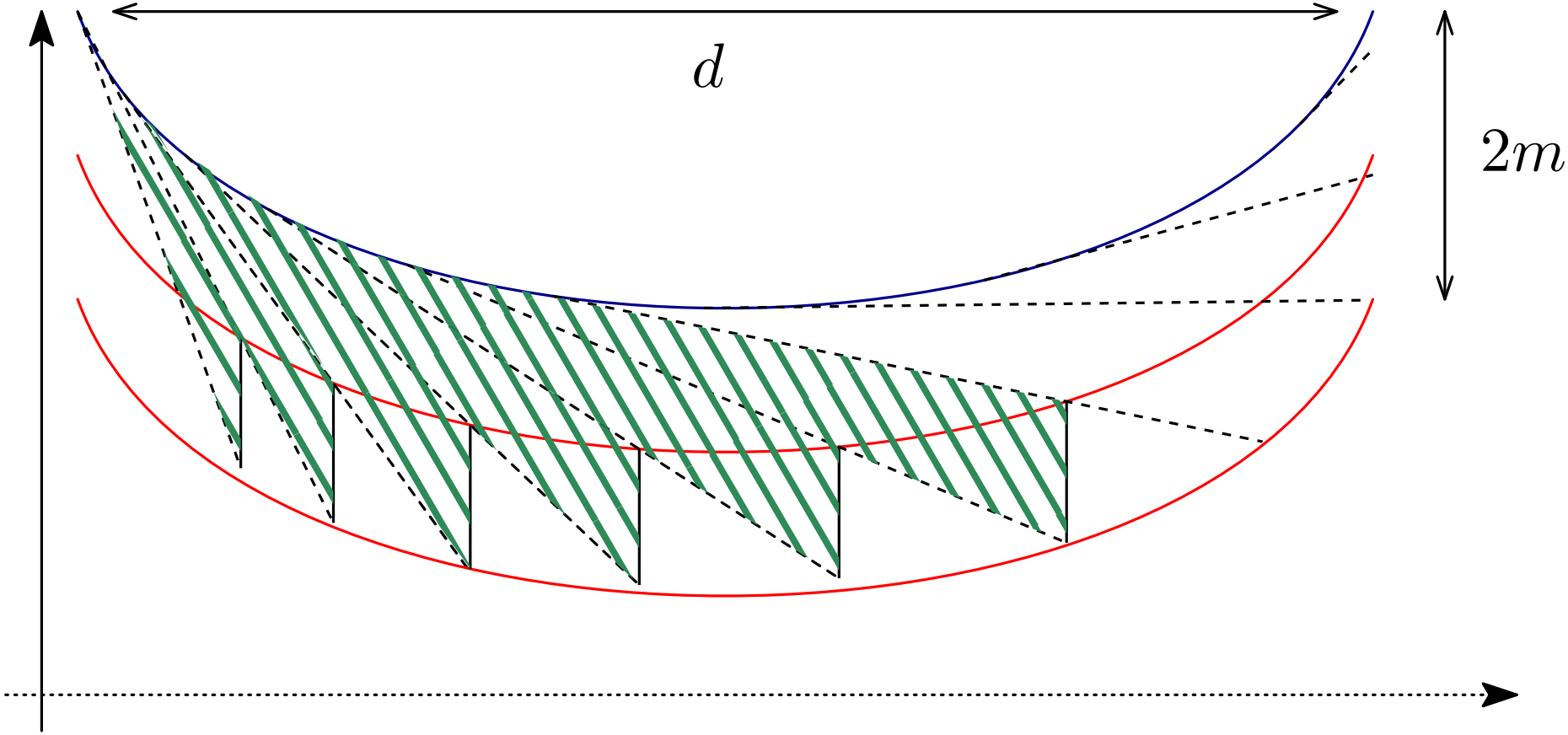


U_j piecewise disjoint

Union of U_j in B_{2m}

$$\sum |U_j| < 2dm$$

Proof

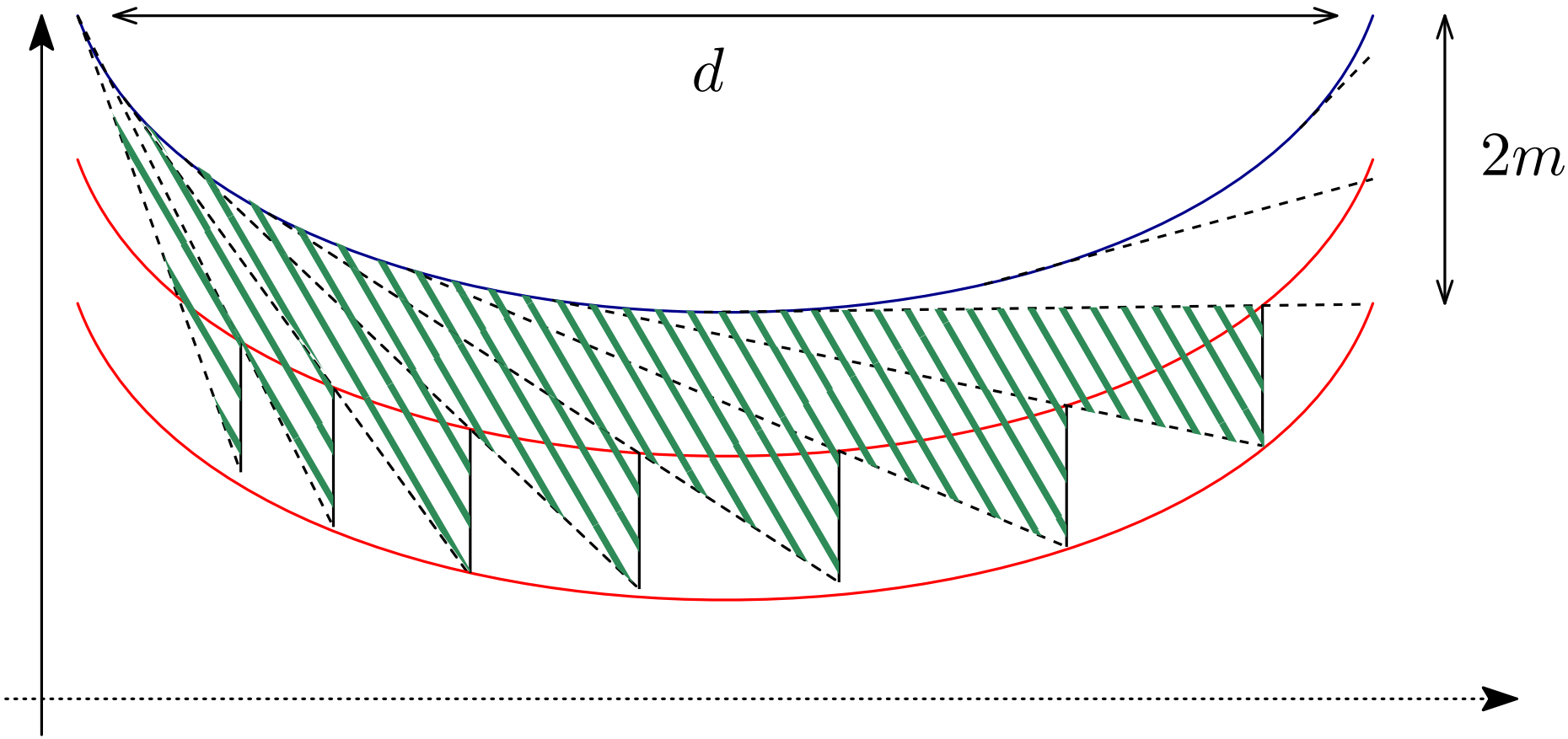


U_j piecewise disjoint

Union of U_j in B_{2m}

$$\sum |U_j| < 2dm$$

Proof

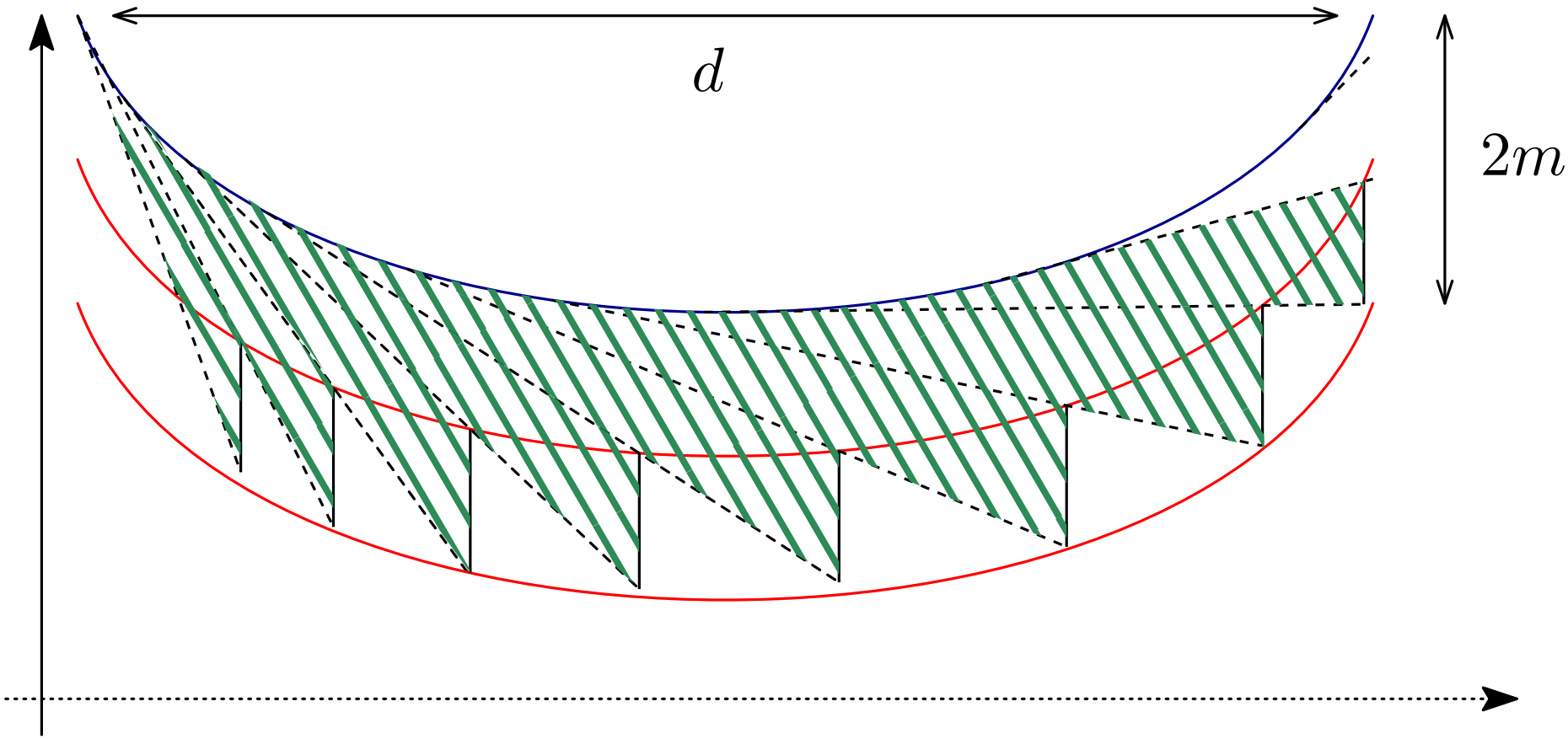


U_j piecewise disjoint

Union of U_j in B_{2m}

$$\sum |U_j| < 2dm$$

Proof

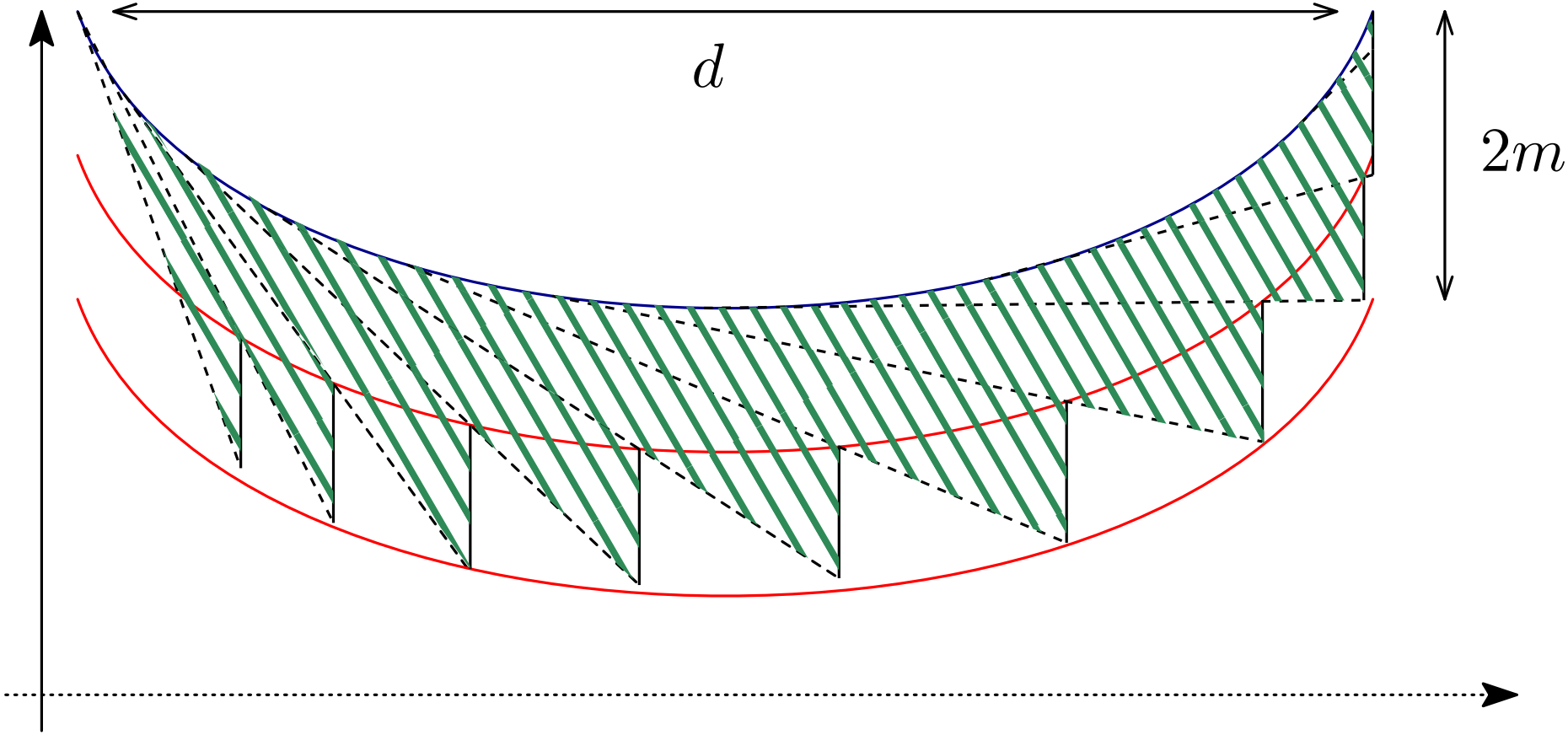


U_j piecewise disjoint

Union of U_j in B_{2m}

$$\sum |U_j| < 2dm$$

Proof



U_j piecewise disjoint

Union of U_j in B_{2m}

$$\sum |U_j| < 2dm$$

Proof

$$|L_j| + |U_j| > \frac{k_j m}{4}$$

$$\sum |L_j| < 2dm$$

$$\sum |U_j| < 2dm$$

Main lemma

$$\sum \frac{k_j m}{4} \leq \sum |L_j| + \sum |U_j| \leq 4dm$$

Benchmark

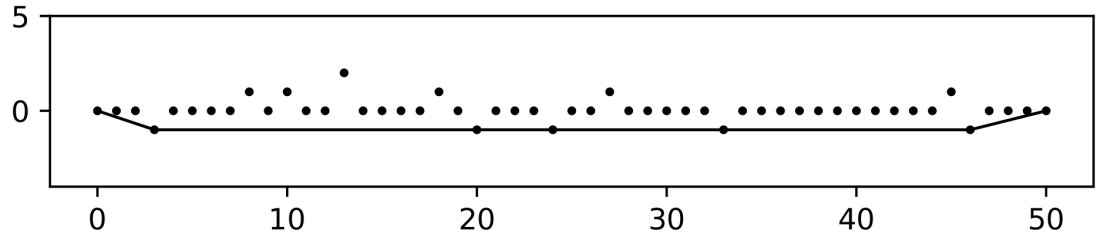
Three families of polynomials

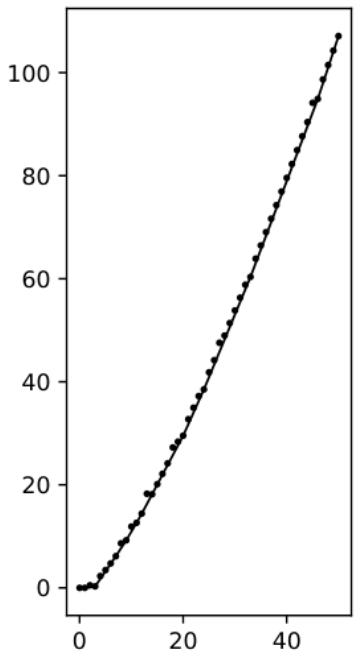
c_j are iid random Gaussian variable with mean 0

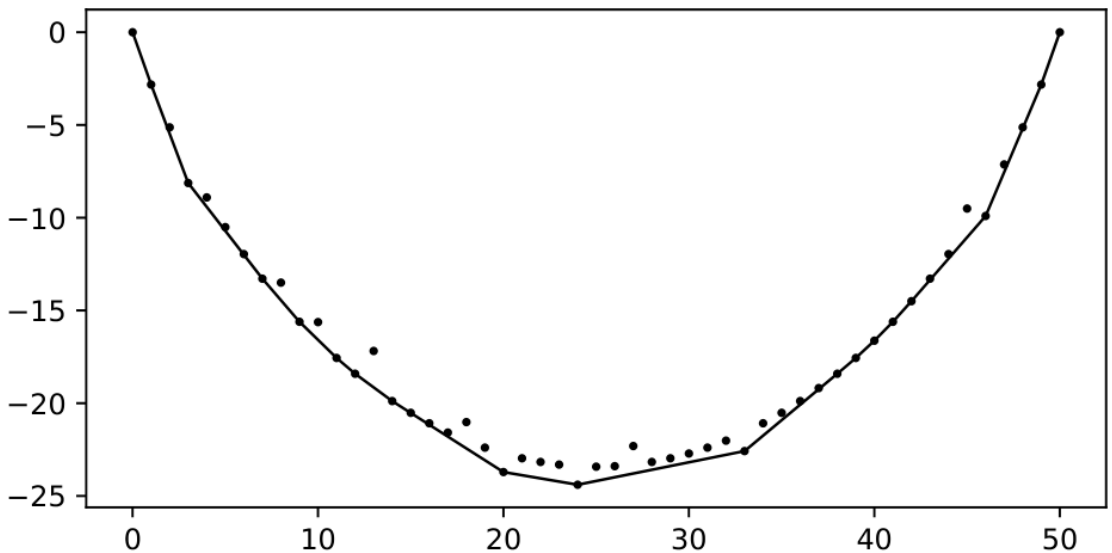
ic: $\sum c_j z^j$

ic: $\sum c_j \sqrt{\binom{d}{j}} z^j$

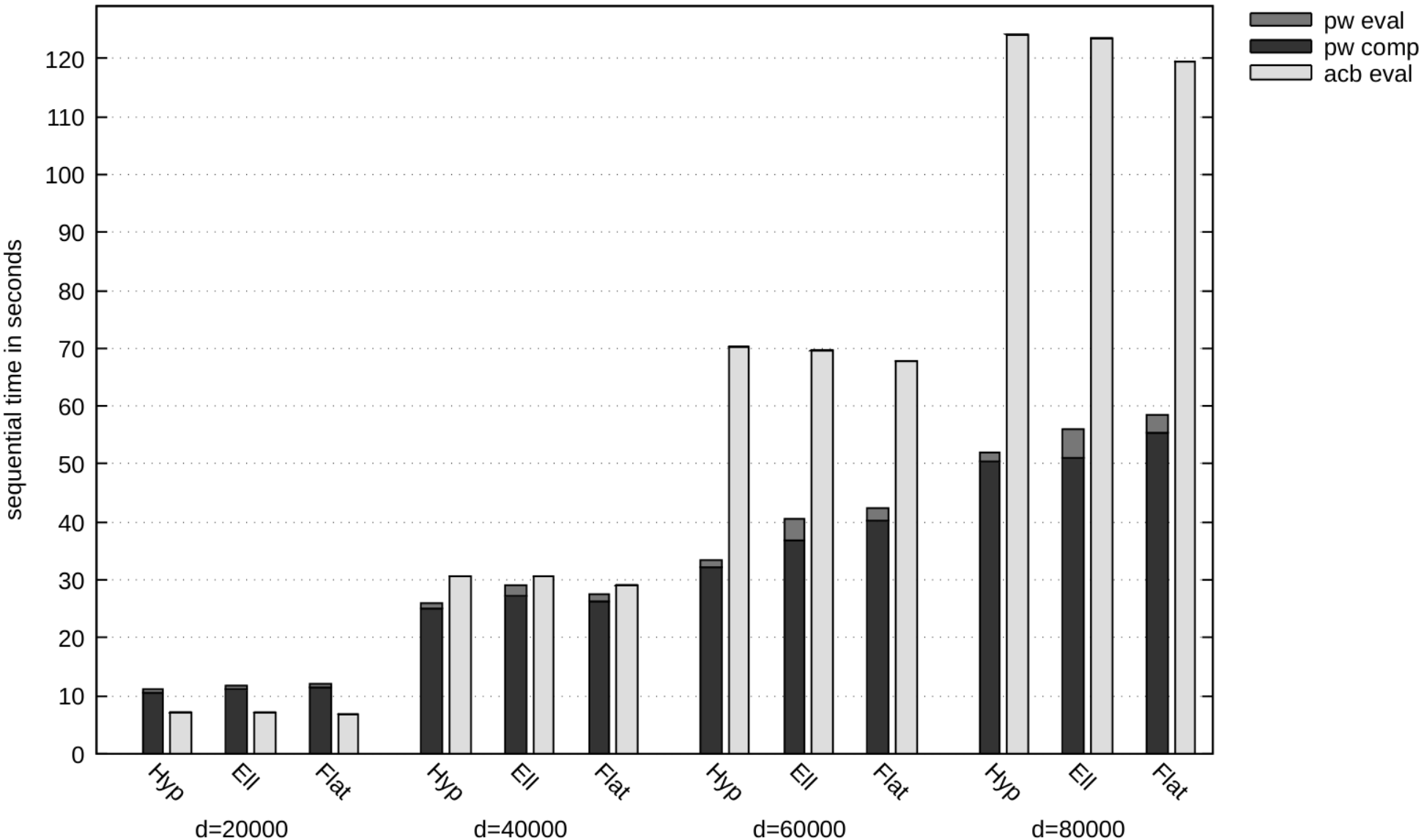
at: $\sum c_j \sqrt{\frac{1}{j!}} z^j$



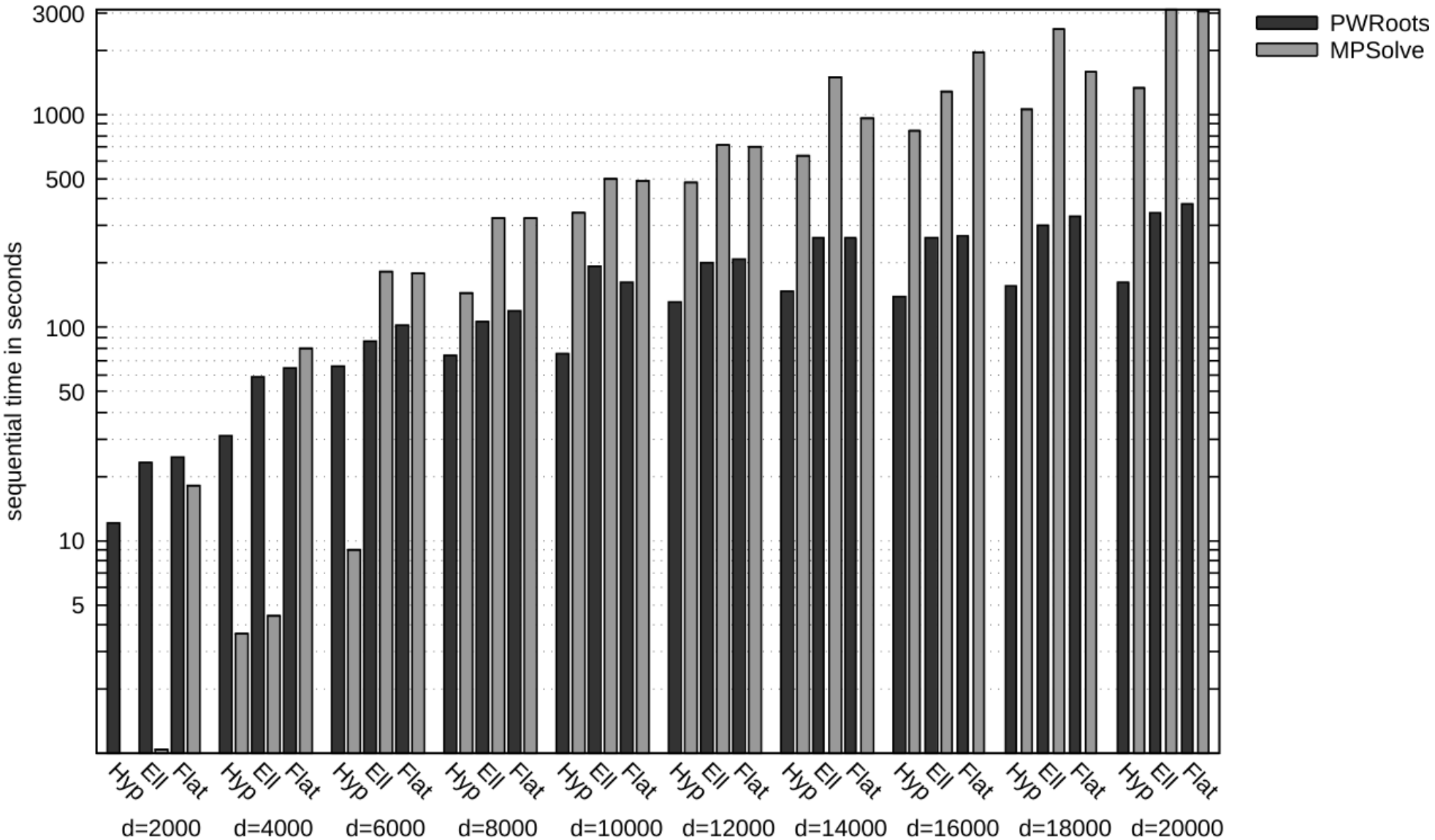




Benchmark


pw eval
pw comp
acb eval
sequential time in seconds
0
10
20
30
40
50
60
70
80
90
100
110
120
Hyp
Ell
Flat
d=20000
d=40000
d=60000
d=80000

Benchmark


PWRoots
MPSolve
sequential time in seconds
3000
1000
500
100
50
10
5
Hyp
Ell
Flat
d=2000
d=4000
d=6000
d=8000
d=10000
d=12000
d=14000
d=16000
d=18000
d=20000

Conclusion

Perspectives

- Bivariate polynomials
- Non integer exponents
- Laplace transform

Thank you!