ON THE EXIT TIME OF A RANDOM WALK WITH POSITIVE DRIFT

Michael Drmota *

joint work with Wojciech Szpankowski

Institut für Diskrete Mathematik und Geometrie

Technische Universität Wien

michael.drmota@tuwien.ac.at

www.dmg.tuwien.ac.at/drmota/

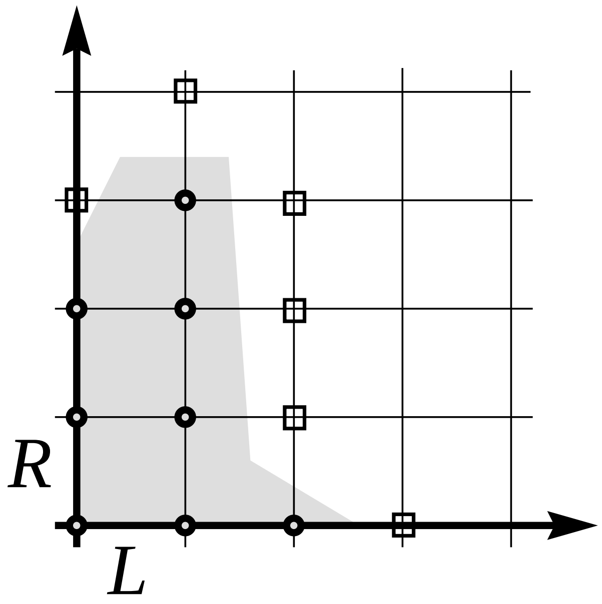
* supported by the Austrian Science Foundation FWF, grant S9604.

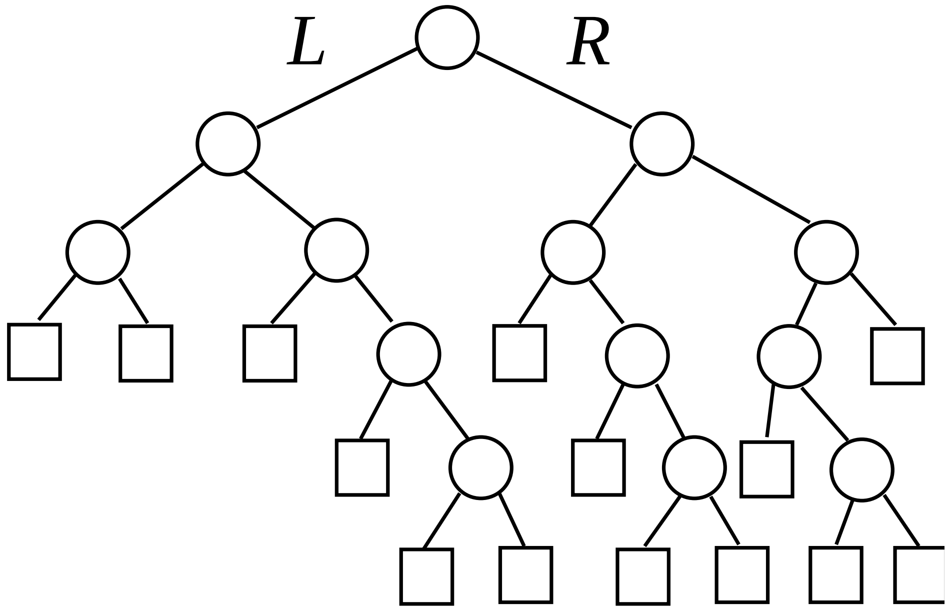
Summary

- **Lattice Paths and Binary Trees**
- **Tunstall-Codes**
- **Number of Paths**
- **Exit Time**
- **Asymptotic Analysis**

Lattice Paths and Binary Trees

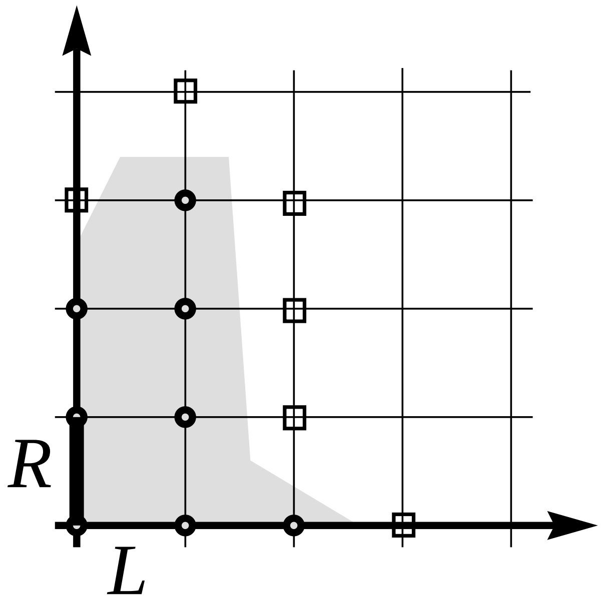
C ... region in the plane

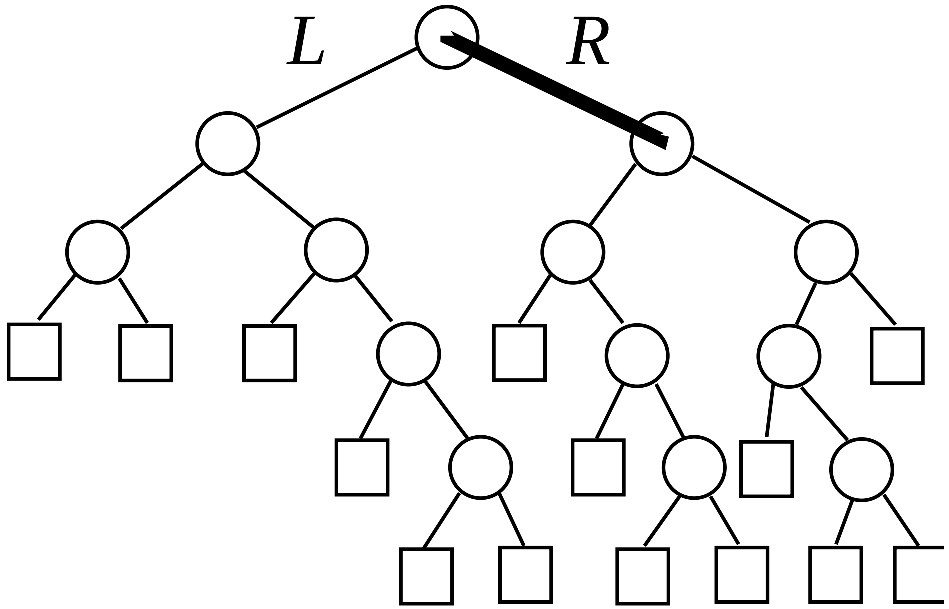




Lattice Paths and Binary Trees

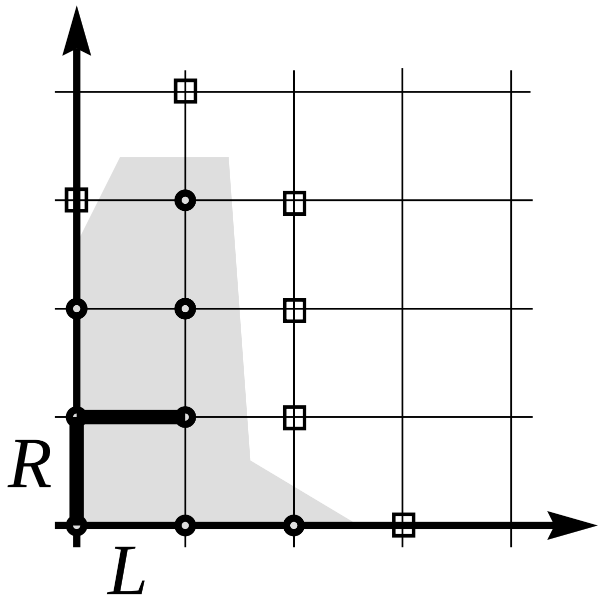
C ... region in the plane

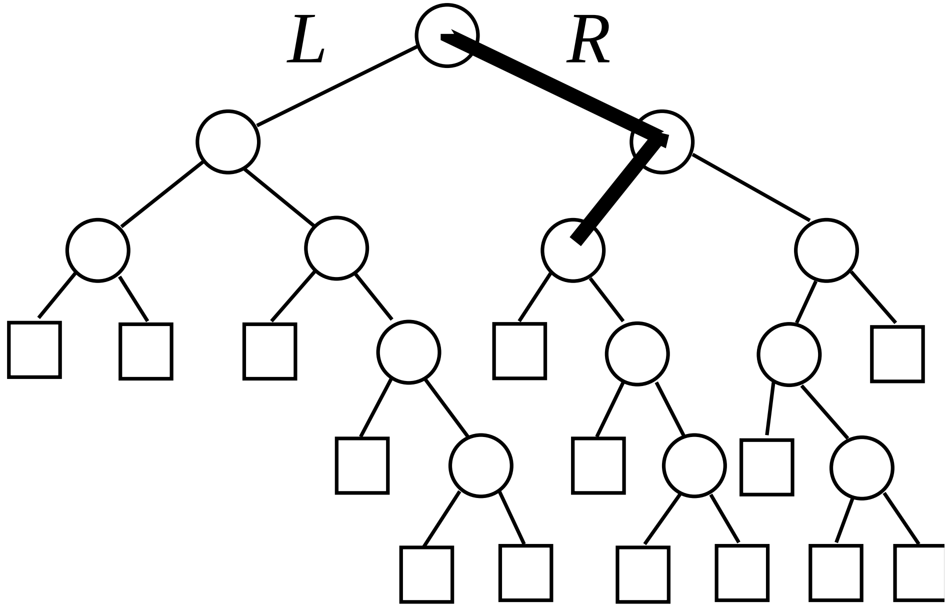




Lattice Paths and Binary Trees

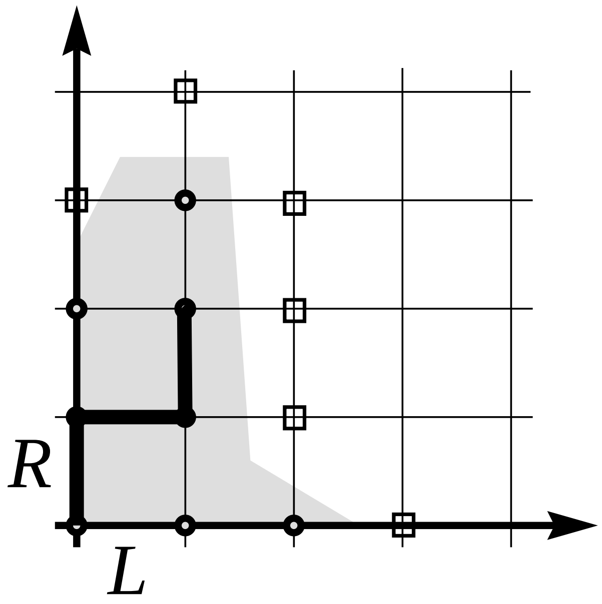
C ... region in the plane

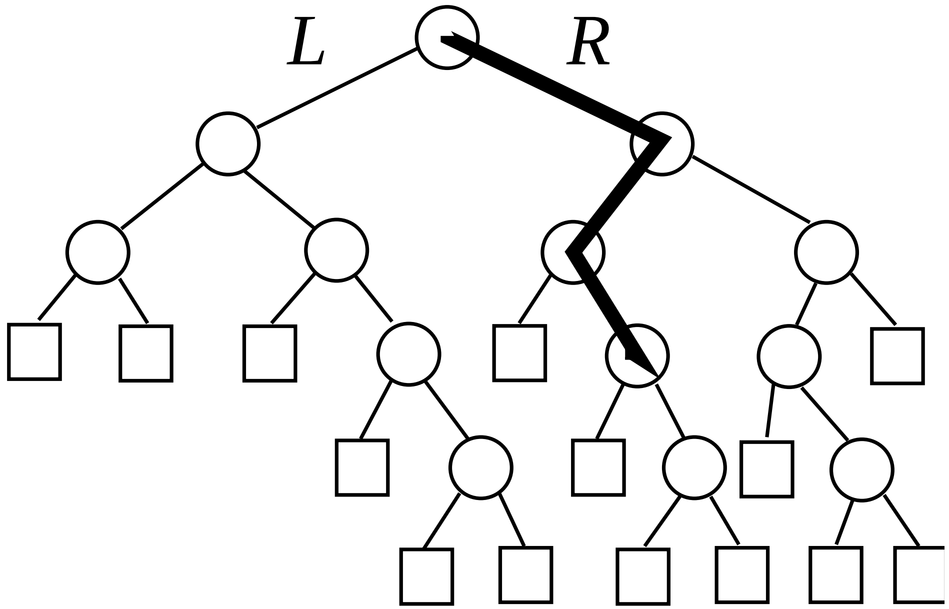




Lattice Paths and Binary Trees

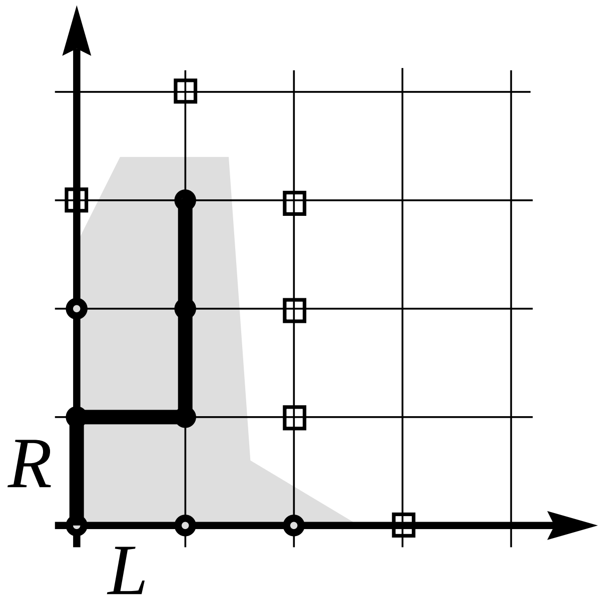
C ... region in the plane

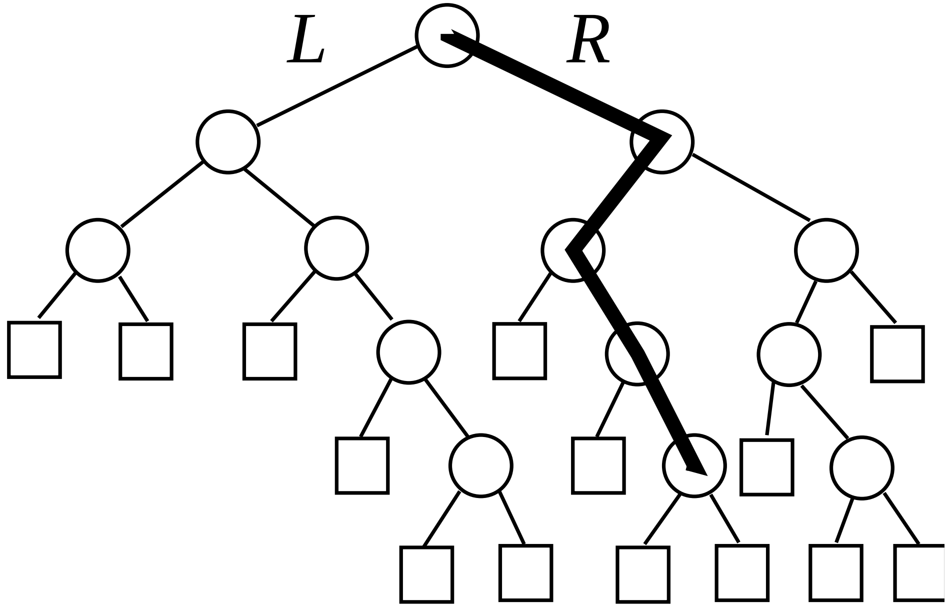




Lattice Paths and Binary Trees

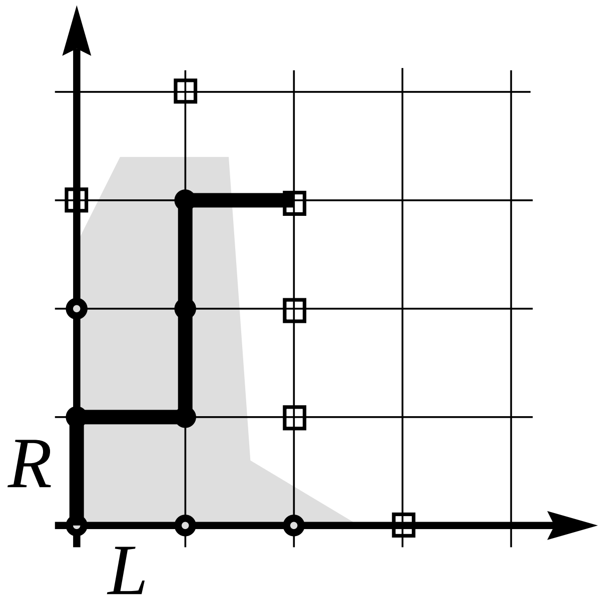
C ... region in the plane

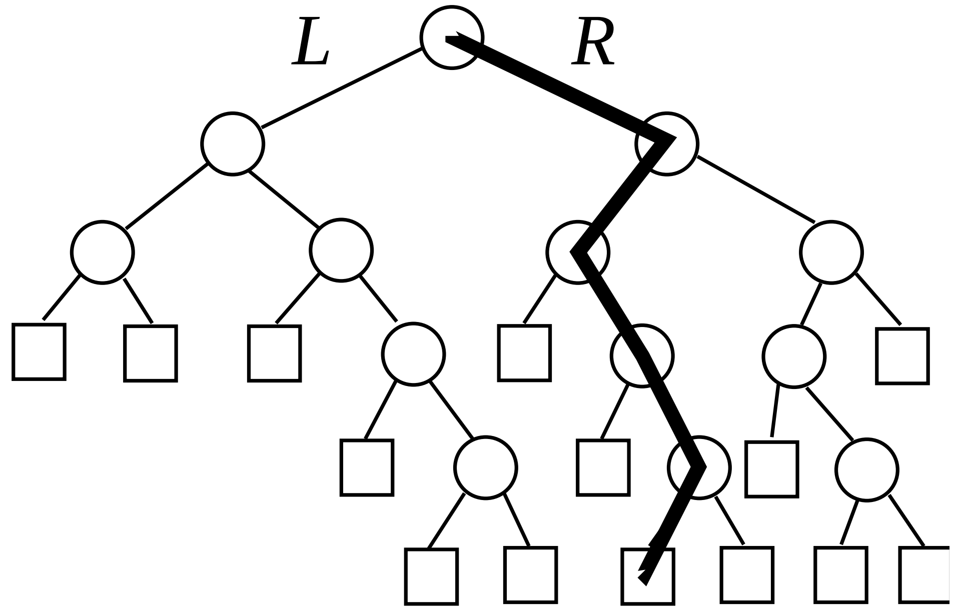




Lattice Paths and Binary Trees

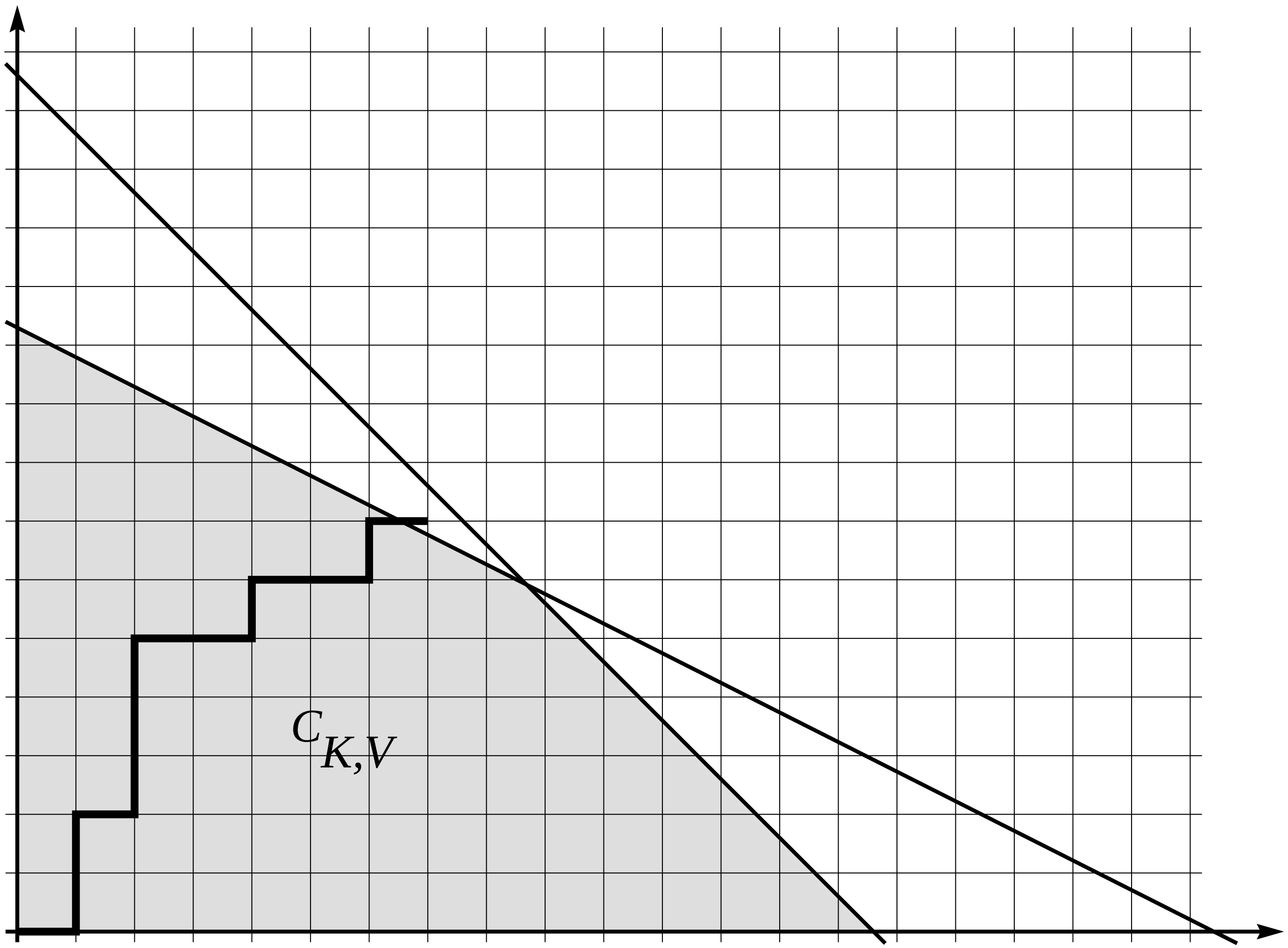
C … region in the plane





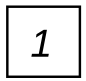
Lattice Paths and Binary Trees

$C_{K,V}$ ... $x_1 \geq 0$, $x_2 \geq 0$, $\boxed{x_1 + x_2 \leq K}$, $\boxed{ax_1 + bx_2 \leq V}$
(W.l.o.g. $a = \log \frac{1}{p}, b = \log \frac{1}{q}$, $p + q = 1$)



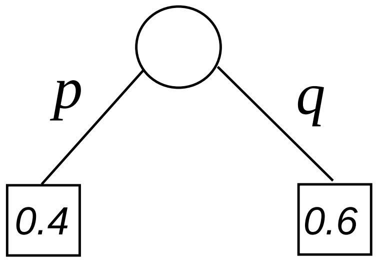
Tunstall-Codes

$p = 0.4$, $q = 0.6$.



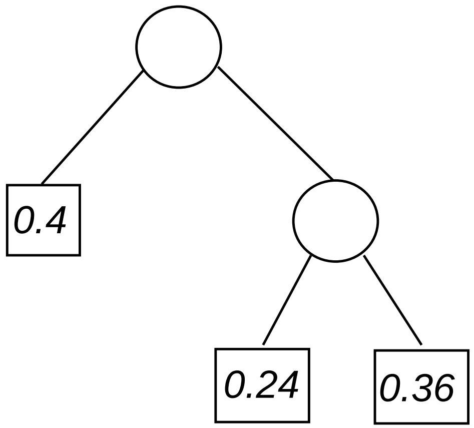
Tunstall-Codes

$p = 0.4$, $q = 0.6$.



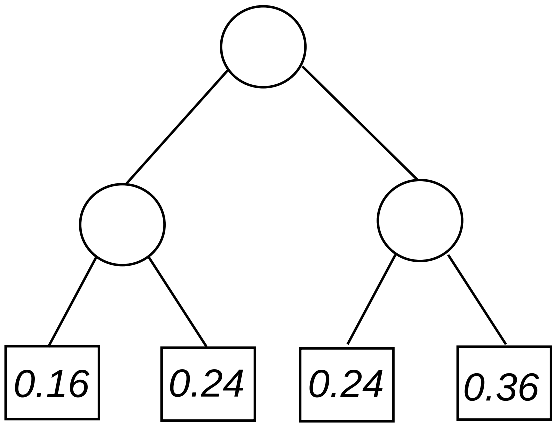
Tunstall-Codes

$p = 0.4,\ q = 0.6.$



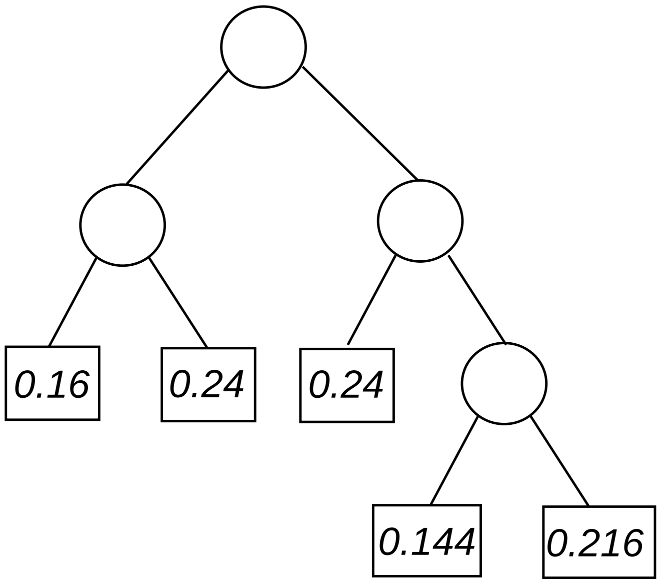
Tunstall-Codes

$p = 0.4,\ q = 0.6.$



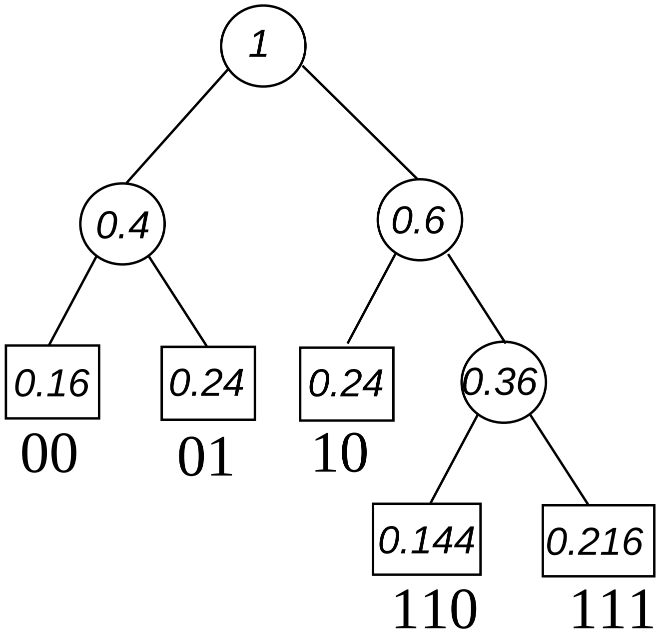
Tunstall-Codes

$p = 0.4,\ q = 0.6.$



Tunstall-Codes

$p = 0.4,\ q = 0.6.$



Tunstall-Codes

y ... node of tree of type (k_1, k_2) [i.e. k_1 times L, k_2 times R]

$$\boxed{P(y) = p^{k_1} q^{k_2}}$$

internal nodes: $P(y) = p^{k_1} q^{k_2} \geq 0.36 \iff \boxed{k_1 \log \frac{1}{p} + k_2 \log \frac{1}{q} \leq \log \frac{1}{0.36}}$

depths: $\boxed{k_1 + k_2 \leq 2}$

external nodes: prefixfree set $= \{00, 01, 10, 110, 111\}$ (Tunstall code)

Number of Paths

K ... integer, V ... positive real number

$$C_{K,V} := \{(x_1,x_2) \in \mathbb{R}^2_{\geq 0} : x_1 + x_2 \leq K,\ x_1 \log_2 \frac{1}{p} + x_2 \log_2 \frac{1}{q} \leq V\}$$

$\mathcal{L}_{K,V}$ ... set of corresponding lattice paths

$T_{K,V}$ ... associated binary tree.

$$|\mathcal{L}_{K,V}| = 1 + \sum_{k_1+k_2 \leq K,\ k_1 \log_2 \frac{1}{p} + k_2 \log_2 \frac{1}{q} \leq V} \binom{k_1+k_2}{k_1}$$

(the number of external nodes = 1 + number of internal nodes)

Number of Paths

Notation

$s_{\mathsf{sp}} = s_{\mathsf{sp}}(K, V)$ is defined by

$$\boxed{\frac{p^{-s_{\mathsf{sp}}} + q^{-s_{\mathsf{sp}}}}{p^{-s_{\mathsf{sp}}} \log \frac{1}{p} + q^{-s_{\mathsf{sp}}} \log \frac{1}{q}} = \frac{K}{V \log 2}}$$

$$s_{\mathsf{sp}} > -1 \iff K/V < \left(p \log_2 \tfrac{1}{p} + 1 \log_2 \tfrac{1}{q}\right)^{-1}.$$

$$T(s) := \frac{p^{-s} \log^2 \frac{1}{p} + q^{-s} \log^2 \frac{1}{q}}{p^{-s} + q^{-s}} - \left(\frac{p^{-s} \log \frac{1}{p} + q^{-s} \log \frac{1}{q}}{p^{-s} + q^{-s}}\right)^2$$

Number of Paths

Notation

$$Q_L(s,x) := \frac{L}{1-e^{sL}} e^{sL\langle \frac{x}{L} \rangle} = \sum_{m\in\mathbf{Z}} \frac{1}{(-s)+\frac{2\pi i m}{L}} e^{\frac{2\pi i m}{L} x},$$

$\langle y \rangle = y - \lfloor y \rfloor$ ... fractional part of a real number y.

$H := p \log(1/p) + q \log(1/q)$ ... **entropy** of the distribution p, q

$H_2 := p \log^2(1/p) + q \log^2(1/q)$.

Number of Paths

Theorem 1 *Suppose that* $\delta > 0$ *is given.*

1. $$\frac{\log 2}{H} \cdot (1+\delta) \le \frac{K}{V} \le \frac{\log 2}{\min\{\log(1/p), \log(1/q)\}} \cdot (1-\delta)$$

$$\frac{\log p}{\log q} \notin \mathbb{Q} \quad \Longrightarrow \quad \boxed{|\mathcal{L}_{K,V}| = \frac{2^V}{H}(1+o(1))}$$

$$\frac{\log p}{\log q} \notin \mathbb{Q} \quad \Longrightarrow \quad \boxed{|\mathcal{L}_{K,V}| = \frac{Q_L(-1, V \log 2)}{H} 2^V + O(2^{V(1-\eta)}),}$$

where $\eta > 0$, *and* $L > 0$ *is the largest real number for which* $\log(1/p)$ *and* $\log(1/q)$ *are integer multiples of* L*:*

$$\frac{\log p}{\log q} = \frac{L \log \frac{1}{p}}{L \log \frac{1}{q}} = \frac{d}{r}.$$

Number of Paths

2. $$\frac{2\log 2}{\log(1/p)+\log(1/q)}\cdot(1+\delta)\le\frac{K}{V}\le\frac{\log 2}{H}\cdot(1-\delta)$$

$$\Longrightarrow$$

$$\boxed{|\mathcal{L}_{K,V}|\sim\sum_{\ell\ge 0}\frac{Q_\delta\left(s_{\mathsf{sp}},(K-\ell)\log p-V\log 2\right)}{(p^{-s_{\mathsf{sp}}}+q^{-s_{\mathsf{sp}}})^{\ell}}\cdot\frac{(p^{-s_{\mathsf{sp}}}+q^{-s_{\mathsf{sp}}})^{K}2^{-V s_{\mathsf{sp}}}}{\sqrt{2\pi K\,T(s_{\mathsf{sp}})}},}$$

where $\delta=\log q-\log p$*.*

Number of Paths

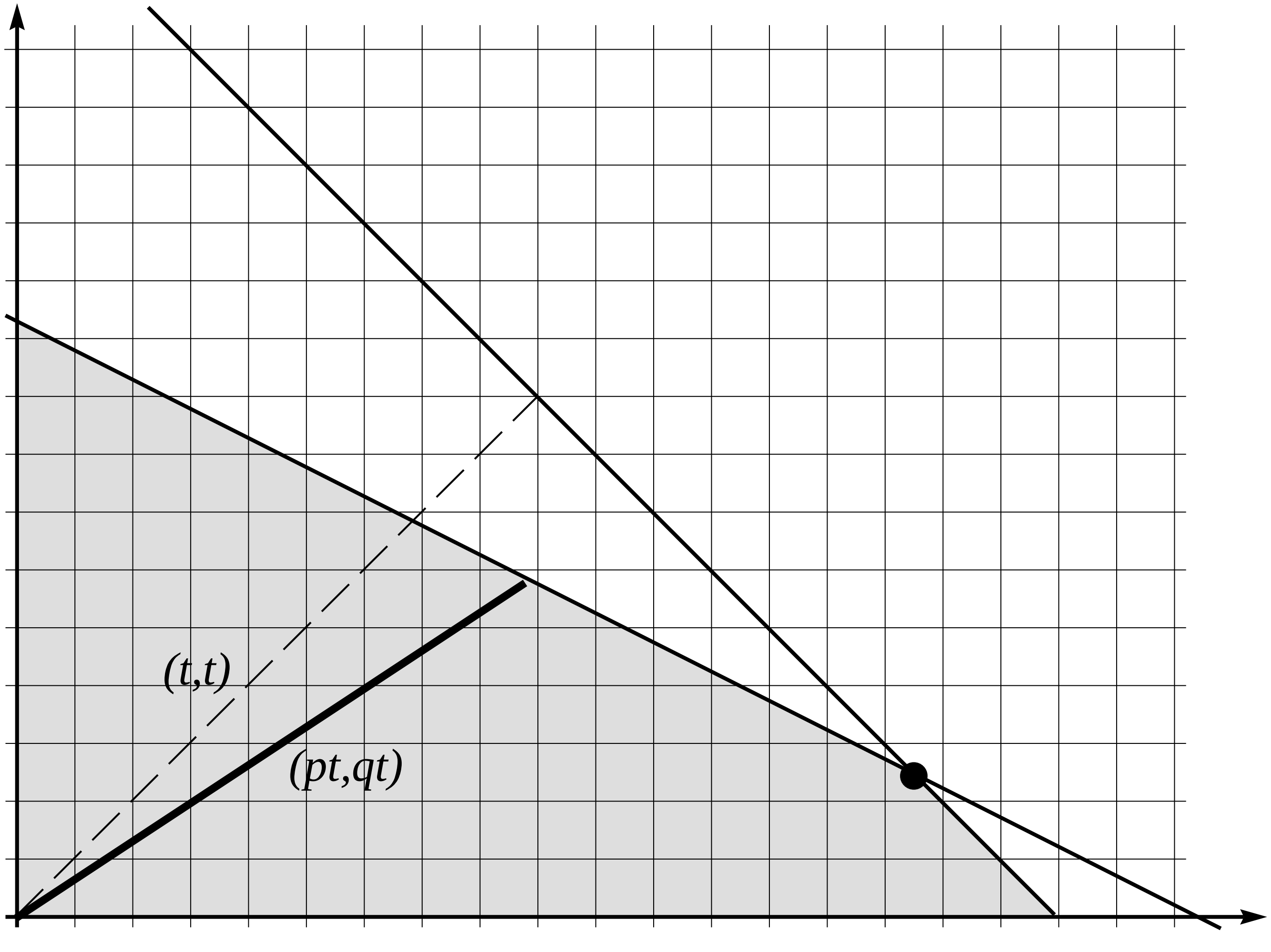
3. $$\frac{\log 2}{\max\{\log(1/p), \log(1/q)\}} \cdot (1+\delta) \le \frac{K}{V} \le \frac{2\log 2}{\log(1/p) + \log(1/q)} \cdot (1-\delta)$$

$$\Longrightarrow \quad \boxed{|\mathcal{L}_{K,V}| = 2^{K+1} - O(2^{K(1-\eta)})}$$

(for some $\eta > 0$).

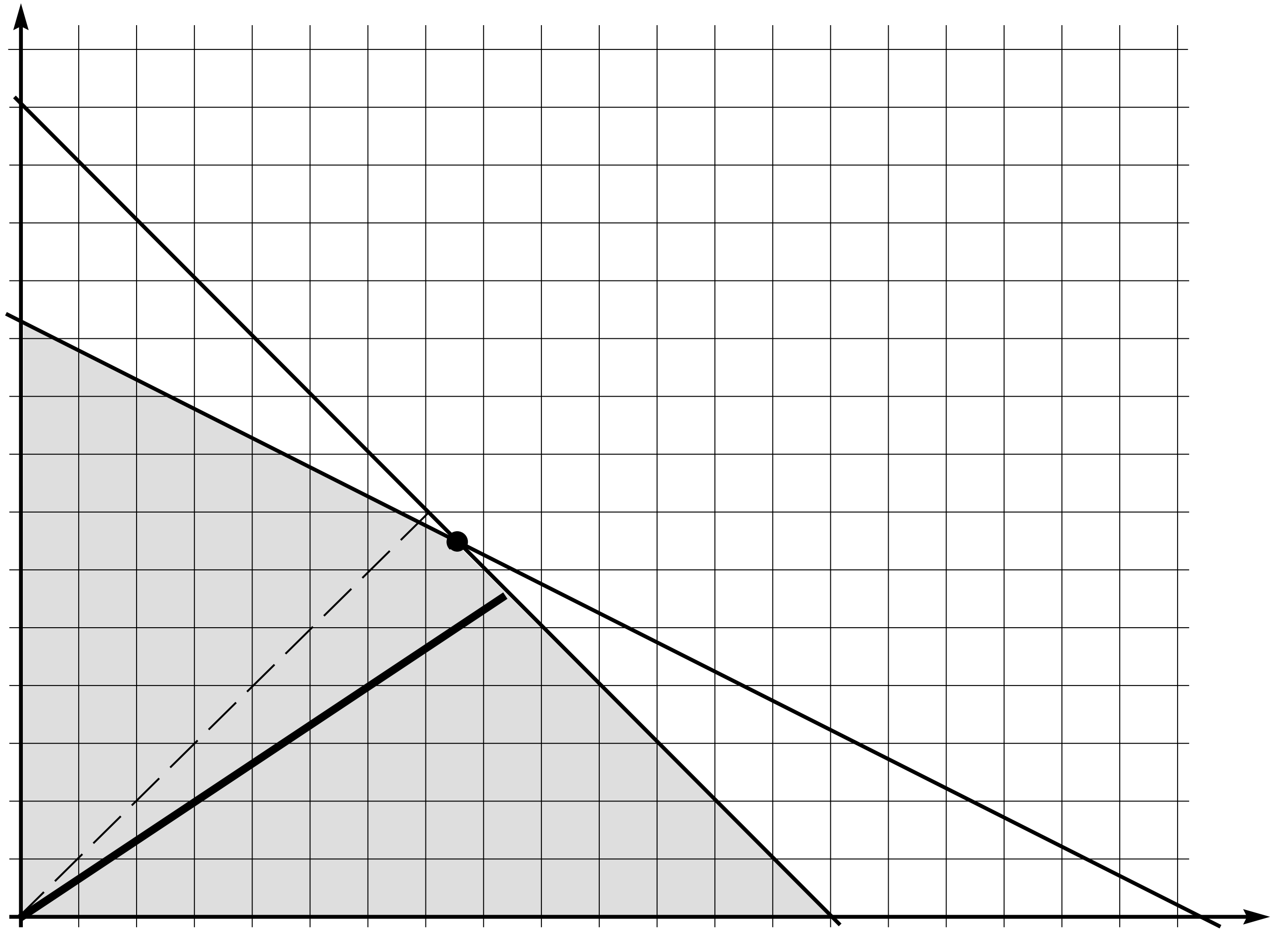
Number of Paths

Case 1



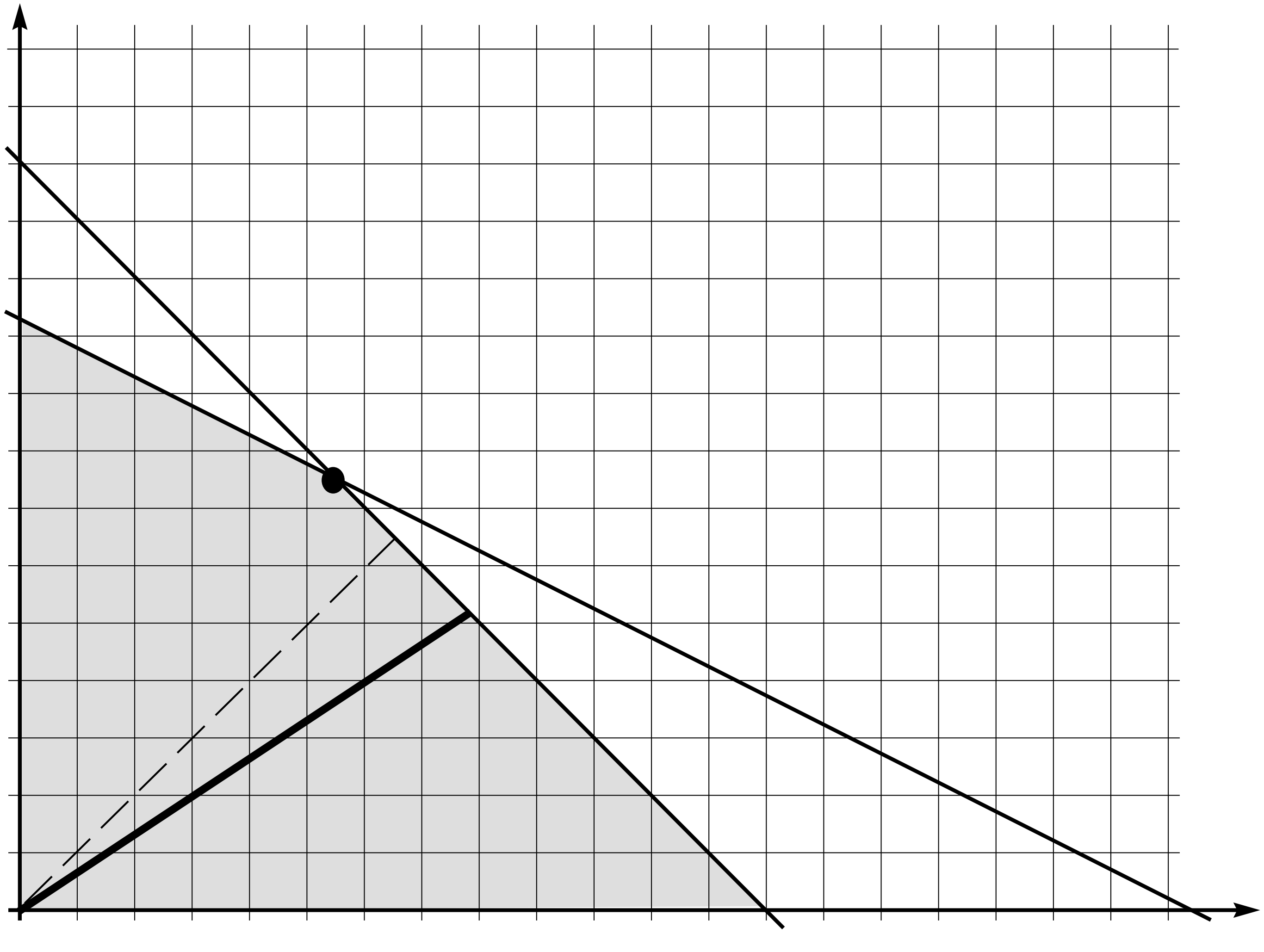
Number of Paths

Case 2



Number of Paths

Case 3



Exit Time

y ... lattice path in $\mathcal{L}_{K,V}$ with k_1 steps R and k_2 steps L

$P(y) := p^{k_1}q^{k_2}$ ... "natural probability distribution"

$\sum_{y\in\mathcal{L}_{K,V}} P(y) = 1$ ... every path y eventually leaves $C_{K,V}$

$D_{K,V}$ ... exit time of this random walk: $|y| = k_1 + k_2$

Equivalently, $D_{K,V}$ is the depth of the external nodes of the corresponding tree (with distribution P).

Exit Time

Theorem 2 *Suppose that $\delta > 0$ is given.*

1. $$\frac{\log 2}{H}\cdot(1+\delta) \le \frac{K}{V} \le \frac{\log 2}{\min\{\log(1/p), \log(1/q)\}}\cdot(1-\delta)$$

$$\Longrightarrow \quad \boxed{\frac{D_{K,V} - \frac{1}{H}\log|\mathcal{L}_{K,V}|}{\left(\left(\frac{H_2}{H^3}-\frac{1}{H}\right)\log|\mathcal{L}_{K,V}|\right)^{1/2}} \xrightarrow{\text{d}} N(0,1)}$$

$$\mathbb{E}\, D_{K,V} = \frac{\log|\mathcal{L}_{K,V}|}{H} + \frac{\log H}{H} + \frac{H_2}{2H^2} + \frac{-\log L + \log(1-e^{-L}) + \frac{L}{2}}{H} + o\left(\frac{1}{\log|\mathcal{L}_{K}}\right.$$

$$\mathbb{V}\, D_{K,V} = \left(\frac{H_2}{H^3} - \frac{1}{H}\right)\log|\mathcal{L}_{K,V}| + O(1).$$

$L = 0$ if $\log p/\log q$ is irrational and $L > 0$ is the largest real number for which $\log(1/p)$ and $\log(1/q)$ are integer multiples of L if $\log p/\log q$ is rational.

2. $$\frac{\log 2}{\max\{\log(1/p), \log(1/q)\}} \cdot (1+\delta) \le \frac{K}{V} \le \frac{\log 2}{H} \cdot (1-\delta)$$

$\Longrightarrow$ *the distribution of* $D_{K,V}$ *is asymptotically concentrated at* $K+1$*:*

$$\boxed{\mathbf{Pr}\{D_{K,V} \neq K+1\} = O(e^{-\eta K})}$$

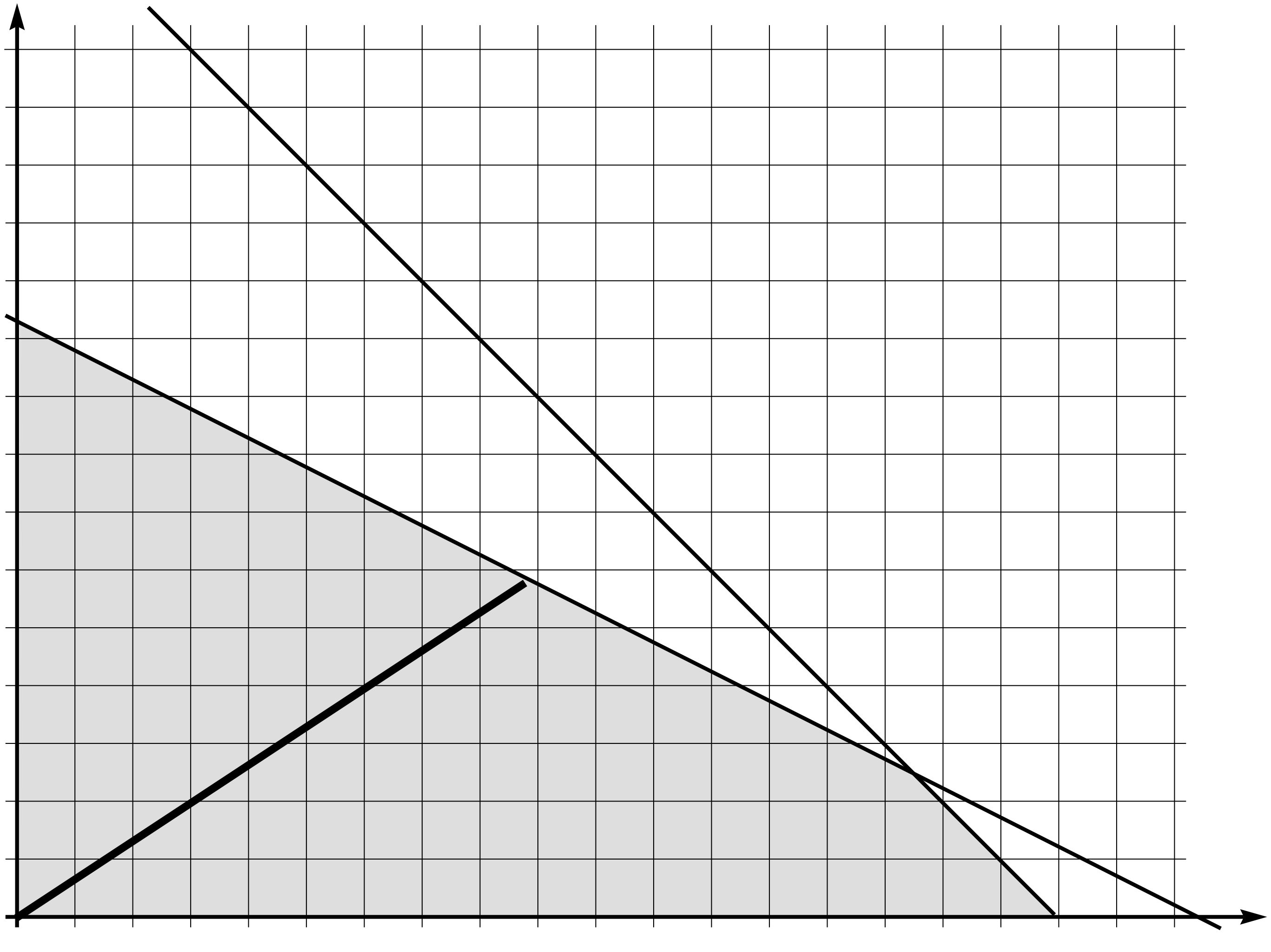
for some $\eta > 0$.

$\mathbb{E}\, D_{K,V} = K + 1 + O(e^{-\eta K})$, $\mathbb{V}\, D_{K,V} = O(e^{-\eta K})$.

Remark. Case 2 of Theorem 2 corresponds to cases 2 and 3 of Theorem 1.

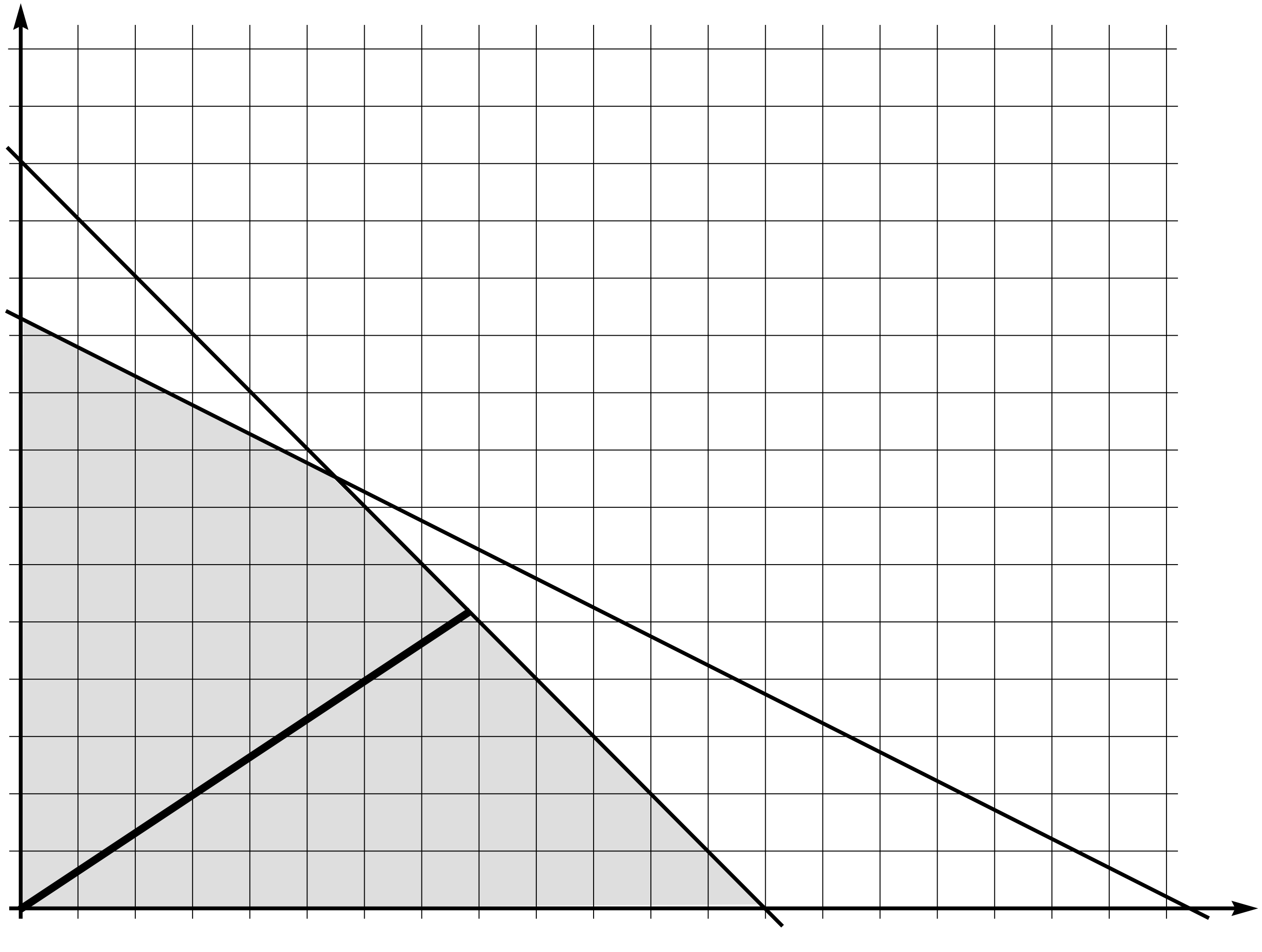
Exit Time

Case 1



Exit Time

Case 2



Asymptotic Analysis

Internal versus external nodes

Lemma 1
T ... m-ary tree,
X ... leaves (=external nodes) of T,
Y ... internal nodes of T

$P(y) := p_1^{k_1} \cdots p_m^{k_m}$ if y is of type $(k_1, \ldots, k_m)$

$$\Longrightarrow \quad \boxed{\sum_{x \in X} P(x) z^{|x|} = (z-1) \sum_{y \in Y} P(y) z^{|y|} + 1}$$

Remark. expected (external) depth $= \sum_{x \in X} |x| P(x) = \sum_{y \in Y} P(y)$.

Asymptotic Analysis

Distribution of zeroes

Lemma 2 (Jacquet, Schachinger) $0 < p < q < 1$
$Z := \{s \in \mathbb{C} : p^{-s} + q^{-s} = 1\}$

(i) $-1 \leq \Re(s) \leq \sigma_0$ *for all* $s \in Z$
($\sigma_0 > 0, 1 + q^{-\sigma_0} = p^{-\sigma_0}$*)*

(ii) *For every integer* k *there uniquely exists* $s_k \in Z$ *with*
$(2k-1)\pi/\log p < \Im(s_k) < (2k+1)\pi/\log p$.
$Z = \{s_k : k \in \mathbb{Z}\}$.

(iii) $\log p/\log q \notin \mathbb{Q} \Longrightarrow s_0 = -1$ *and* $\Re(s_k) > -1$ *for all* $k \neq 0$.

(iv) $\log p/\log q = d/r \in \mathbb{Q}$, $\gcd(r,d) = 1 \Longrightarrow$
$\Re(s_k) = -1$ *if and only if* $k \equiv 0 \bmod d$, *further*

$$s_k = s_{k \bmod d} + (k - (k \bmod d))\frac{2\pi i}{\log p}.$$

Asymptotic Analysis

Recall: $k_1 \log_2 \frac{1}{p} + k_2 \log_2 \frac{1}{q} \leq V \iff P(y) \geq 1/v \quad (v = 2^V)$

$$\boxed{A_K(v) := \sum_{y: P(y) \geq 1/v,\, |y| \leq K} 1}$$

is the number of lattice paths with endpoints **contained in** $C_{K,V}$

$$\Longrightarrow \quad \boxed{|\mathcal{L}_{K,V}| = A_K(v) + 1 = A_K(2^V) + 1}$$

(the number of external nodes = 1 + number of internal nodes)

Asymptotic Analysis

$$\boxed{S_K(v,z) := \sum_{y: P(y) \geq 1/v,\, |y| \leq K} P(y) z^{|y|}}$$

is the "probability generating function of the internal nodes"

$$D_K(v,z) := \mathbb{E}\, z^{D_{K,V}} \qquad (v = 2^V)$$

$$\Longrightarrow \quad \boxed{D_K(v,z) = (z-1) S_K(v,z) + 1} \qquad (v \geq 1)$$

Asymptotic Analysis

Recurrences

$$A_K(v) = 0 \quad (v < 1, K \geq 0)$$

$$A_{K+1}(v) = 1 + A_K(vp) + A_K(vq)$$

$$S_K(v,z) = 0 \quad (v < 1, K \geq 0)$$

$$S_{K+1}(v,z) = 1 + pzS_K(vp,z) + qzS_K(vq,z).$$

Asymptotic Analysis

Mellin transform $\quad f^*(s) = \int_0^\infty f(v) v^{s-1} dv$

recurrence $\Longrightarrow$

$$A^*_{K+1}(s) = -\frac{1}{s} + (p^{-s} + q^{-s}) A^*_K(s) \qquad (\Re(s) < -1)$$

$$\Longrightarrow \boxed{A^*_K(s) = -\frac{1 - (p^{-s} + q^{-s})^{K+1}}{s(1 - (p^{-s} + q^{-s}))}}$$

$$S^*_{K+1}(s,z) = -\frac{1}{s} + (zp^{1-s} + zq^{1-s}) S^*_K(s,z) \qquad (\Re(s) < 0).$$

$$\Longrightarrow \boxed{S^*_K(s,z) = -\frac{1 - (z(p^{1-s} + q^{1-s}))^{K+1}}{s(1 - z(p^{1-s} + q^{1-s}))}.}$$

Asymptotic Analysis

Inverse Mellin transform

$$A_K(v) = \frac{1}{2\pi i} \lim_{T\to\infty} \int_{\sigma-iT}^{\sigma+iT} A_K^*(s) v^{-s}\, ds \quad (\sigma < -1)$$

$$S_K(v,z) = \frac{1}{2\pi i} \lim_{T\to\infty} \int_{\sigma-iT}^{\sigma+iT} S_K^*(s,z) v^{-s}\, ds \quad (\sigma < 0)$$

Asymptotic Analysis

$$A_K(v) = -\frac{1}{2\pi i} \lim_{T\to\infty} \int_{\sigma-iT}^{\sigma+iT} \frac{1-(p^{-s}+q^{-s})^{K+1}}{s(1-(p^{-s}+q^{-s}))} v^{-s}\, ds$$

Case 1: $\frac{1}{s(1-(p^{-s}+q^{-s}))}$ dominates, $s_{\text{sp}} < -1$

Case 2: $\frac{(p^{-s}+q^{-s})^{K+1}}{s(1-(p^{-s}+q^{-s}))}$ dominates, $-1 < s_{\text{sp}} < 0$

Case 3: $\frac{1}{s}$ dominates, $s_{\text{sp}} > 0$.

PROBLEM: No absolute convergence!

Asymptotic Analysis

Case 1

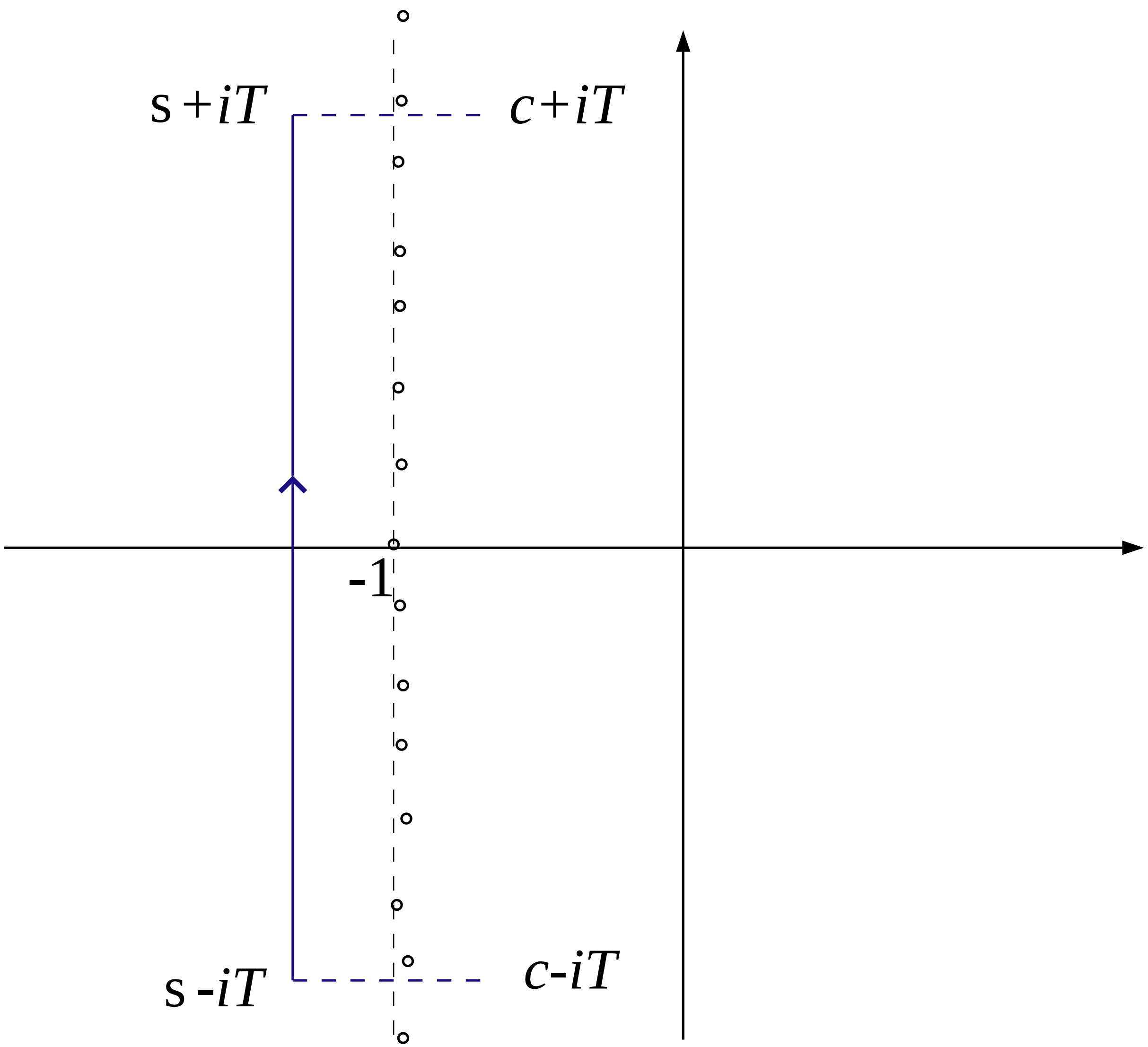
$A_k(v) = I_1 + I_2$:

$$I_1 = -\frac{1}{2\pi i} \lim_{T\to\infty} \int_{\sigma-iT}^{\sigma+iT} -\frac{1}{s(1-(p^{-s}+q^{-s}))} v^{-s}\, ds,$$
$$I_2 = -\frac{1}{2\pi i} \lim_{T\to\infty} \int_{\sigma-iT}^{\sigma+iT} \frac{(p^{-s}+q^{-s})^{K+1}}{s(1-(p^{-s}+q^{-s}))} v^{-s}\, ds.$$

s_{sp} = saddle point of $(p^{-s}+q^{-s})^K v^{-s}$
$s_{\mathrm{sp}} < -1 \implies I_2$ is negligible

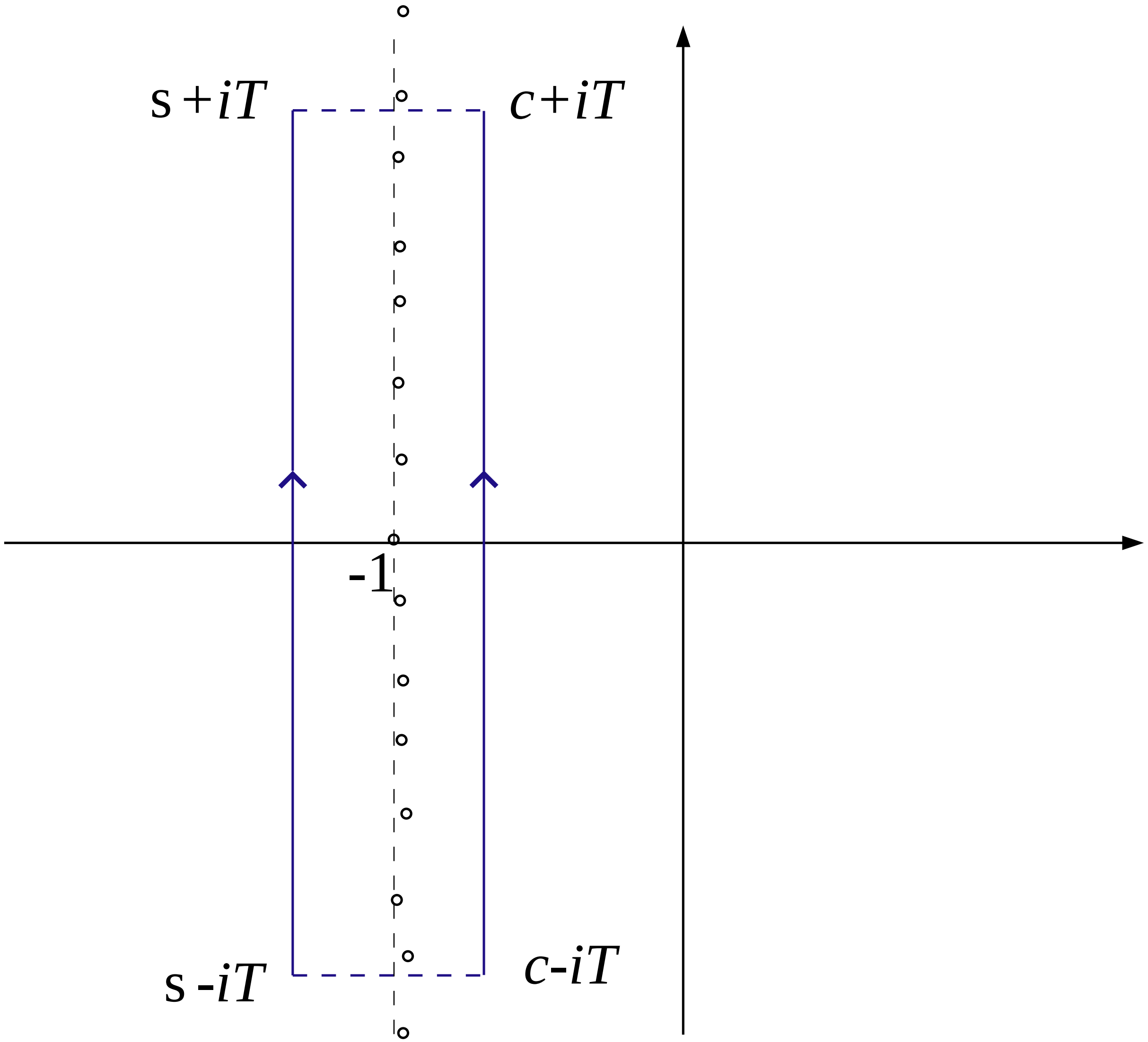
Asymptotic Analysis

Shift of Integration for I_1



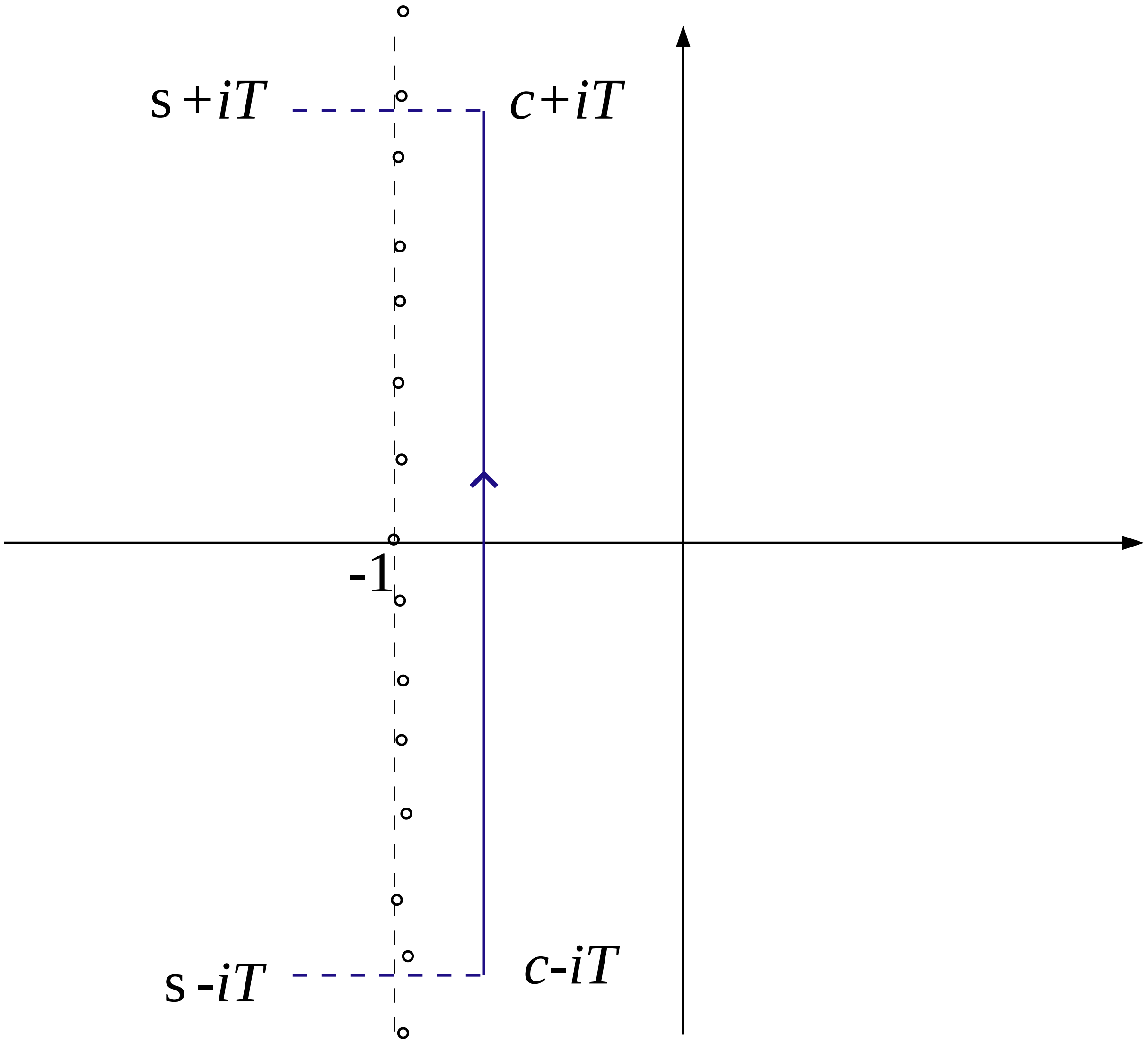
Asymptotic Analysis

Shift of Integration for I_1



Asymptotic Analysis

Shift of Integration for I_1



Asymptotic Analysis

Case 1

$$
\begin{aligned}
I_1 &= -\lim_{T\to\infty} \sum_{s'\in Z,\ \Re(s')=-1, |\Im(s')|<T} \mathrm{Res}(A^*(s)\, v^{-s},\ s = s') \\
&\quad -\frac{1}{2\pi i} \lim_{T\to\infty} \int_{\sigma-iT}^{\sigma+iT} \frac{1}{s(1-p^{-s}-q^{-s})} v^{-s}\, ds \\
&= -\lim_{T\to\infty} \sum_{s'\in Z,\ \Re(s')=-1, |\Im(s')|<T} \frac{v^{-s'}}{s' H(s')} \\
&\quad -\frac{1}{2\pi i} \lim_{T\to\infty} \int_{\sigma-iT}^{\sigma+iT} \frac{1}{s(1-p^{-s}-q^{-s})} v^{-s}\, ds
\end{aligned}
$$

$H(s) = -p^{-s} \log p - q^{-s} \log q.$

PROBLEM: No absolute convergence!

Asymptotic Analysis

Case 1

$\log p / \log q \notin \mathbb{Q} \Longrightarrow$

$$I_1 = \frac{v}{H}(1 + o(1)) = \frac{2^V}{H}(1 + o(1))$$

$\log p / \log q \in \mathbb{Q} \Longrightarrow$

$$I_1 = \sum_{m \in \mathbb{Z}} \frac{v^{1-2\pi i \frac{m}{L}}}{(1 - 2\pi i \frac{m}{L})H} + O(v^{1-\eta}) = 2^V \frac{Q_L(-1, V \log 2)}{H} + O(v^{1-\eta}),$$

Asymptotic Analysis

Case 2

Shift to $\Re s = s_{\mathsf{sp}} \in (-1, 0)$:

$$A_K(v) = -\frac{1}{2\pi i} \lim_{T\to\infty} \int_{s_{\mathsf{sp}}-iT}^{s_{\mathsf{sp}}+iT} \frac{1-(p^{-s}+q^{-s})^{K+1}}{s(1-(p^{-s}+q^{-s}))} v^{-s}\, ds, \tag{1}$$

Infinitely many saddle points:

$$s_h = s_{\mathsf{sp}} + it_h, \quad t_h = 2\pi h/(\log p - \log q).$$

$\Longrightarrow$

$$A_K(v) \sim \sum_{h\in\mathbb{Z}} \frac{1}{\sqrt{2\pi K\, T(s_{\mathsf{sp}})}} \frac{(p^{-s_{\mathsf{sp}}}+q^{-s_{\mathsf{sp}}})^{K+1} 2^{-s_{\mathsf{sp}}V} p^{-it_h(K+1)} 2^{-iVt_h}}{(s_{\mathsf{sp}}+it_h)(1-(p^{-s_{\mathsf{sp}}}+q^{-s_{\mathsf{sp}}})p^{-it_h})}.$$

Asymptotic Analysis

Case 3

Shift to $\Re s = s_{\mathsf{sp}} > 0$:

$$A_K(v) = 2^{K+1} - \frac{1}{2\pi i} \lim_{T\to\infty} \int_{s_{\mathsf{sp}}-iT}^{s_{\mathsf{sp}}+iT} \frac{1-(p^{-s}+q^{-s})^{K+1}}{s(1-(p^{-s}+q^{-s}))} v^{-s}\, ds$$

$$= 2^{K+1} + O(2^{K(1-\eta)}).$$

Thank You!