On approximation of goodness-of-fit statistics for discrete three dimensional data

Assylbekov[1] Zh. A. Ulyanov[2] V.V.
Zubov[2] V. N.

[1] *Graduate School of Science, Hiroshima University*

[2] *Faculty of Computational Mathematics and Cybernetics, Lomonosov Moscow State University*

August 27, 2008

Abstract

We study rate of convergence for approximation of power-divergence statistics $\{T_\lambda(\boldsymbol{Y}), \lambda \in \mathbb{R}\}$, constructed for n observations of a random variable Y with three possible outcomes. We prove that

$$\Pr(T_\lambda(\boldsymbol{Y}) < c) = K_2(c) + O\left(n^{-100/146}(\log n)^{315/146}\right),$$

where $K_2(c)$ is a distribution function of chi-square distribution with 2 degrees of freedom. The proof is based on Huxley (1993) result about approximation of number of lattice points in large convex bodies.

Key words: *approximation, Huxley theorem, curvature, chi-square distribution, power-divergence statistics.*

1 Introduction and main result

Let $\boldsymbol{Y} = (Y_1, Y_2, Y_3)'$ be a random vector with multinomial distribution $M_3(n, \boldsymbol{\pi})$, that is

$$\Pr(Y_1 = n_1, Y_2 = n_2, Y_3 = n_3) = \begin{cases} n! \prod_{j=1}^{3} \left(\pi_j^{n_j}/n_j!\right) & n_j = 0, \ldots, n \ (j = 1, 2, 3) \\ & \text{and } \sum_{j=1}^{3} n_j = n, \\ 0 & \text{otherwise,} \end{cases}$$

where $\boldsymbol{\pi} = (\pi_1, \pi_2, \pi_3)'$, $\pi_j > 0$, $\sum_{j=1}^{3} \pi_j = 1$. We consider a simple hypothesis $\mathrm{H}_0 : \boldsymbol{\pi} = \boldsymbol{p}$ (here $\boldsymbol{p}$ is a fixed vector with non-zero components) under alternative hypothesis $\mathrm{H}_1 : \boldsymbol{\pi} \neq \boldsymbol{p}$. It is often used in this case a test from so-called power-divergence family of statistics. It has a form

$$T_\lambda(\boldsymbol{Y}) = \frac{2}{\lambda(\lambda+1)} \sum_{j=1}^{3} Y_j \left[\left(\frac{Y_j}{np_j} \right)^{\lambda} - 1 \right], \ \lambda \in \mathbb{R},$$

where $\boldsymbol{p} = (p_1, p_2, p_3)'$, $p_j > 0$ $(j = 1, 2, 3)$ and $\sum_{j=1}^{3} p_j = 1$.

R e m a r k 1. If $\lambda = 0$ or $\lambda = -1$ then T_0 and T_{-1} are defined as the limits of T_λ when $\lambda \to 0$ or $\lambda \to -1$ correspondingly.

R e m a r k 2. These statistics were introduced in [1] and [2] and were denoted by $2nI^\lambda(\boldsymbol{Y})$. If $\lambda = 1, \lambda = -1/2$ and $\lambda = 0$ we get Pearson's chi-square test, loglikelihood ratio statistic and Freeman-Tukey statistic correspondingly.

Our aim is to get approximation for $\Pr(T_\lambda(\boldsymbol{Y}) < c)$, where c here and everywhere below is a positive constant. Since the components of $\boldsymbol{Y}$ are connected by identity

$$Y_1 + Y_2 + Y_3 = n,$$

let us consider variables

$$X_j = (Y_j - np_j)/\sqrt{n}, \ j = \overline{1,3}, \ \boldsymbol{X} = (X_1, \, X_2)^T,$$

provided that null hypothesis holds. The components of the vector $\boldsymbol{X}$ are concentrated on the lattice

$$L = \{x = (x_1, \, x_2)^T; \, x = (\boldsymbol{m} - n\boldsymbol{p})/\sqrt{n}, \, \boldsymbol{p} = (p_1, \, p_2)^T, \, \boldsymbol{m} = (n_1, \, n_2)^T\},$$

where n_j are non-negative integers.

We have

$$\Pr\left(T_\lambda(\boldsymbol{Y}) < c\right) = \Pr\left(T_\lambda(X_1, X_2) < c\right) = \Pr\left(\boldsymbol{X} \in B^\lambda)\right),$$

where

$$B^\lambda = \{(x,y) : T_\lambda(x,y) < c\} \tag{1}$$

and

$$\begin{aligned} T_\lambda(x,y) = {} & \frac{2}{\lambda(\lambda+1)}(np_1 + \sqrt{n}x)\left[\left(1 + \frac{x}{\sqrt{n}p_1}\right)^\lambda - 1\right] \\ & + \frac{2}{\lambda(\lambda+1)}(np_2 + \sqrt{n}y)\left[\left(1 + \frac{y}{\sqrt{n}p_2}\right)^\lambda - 1\right] \\ & + \frac{2}{\lambda(\lambda+1)}(np_3 - \sqrt{n}(x+y))\left[\left(1 - \frac{x+y}{\sqrt{n}p_3}\right)^\lambda - 1\right]. \end{aligned} \tag{2}$$

The set B^λ is so-called extended convex set. We prove it in Section 3. Now let us remind

Definition 1. A set $B \subset \mathbb{R}^2$ is called *an extended convex set* when it can be represented in a form:

$$\begin{aligned} B &= \{(x,y) : \lambda_1(y) < x < \theta_1(y),\ y \in B_1\} \\ &= \{(x,y) : \lambda_2(x) < y < \theta_2(x),\ x \in B_2\}. \end{aligned}$$

where $B_1 \subset \mathbb{R}$, $B_2 \subset \mathbb{R}$, and λ_1, θ_1, λ_2, θ_2 are continuous functions in $\mathbb{R}$.

For the random vector $\boldsymbol{X}$ defined above J.Yarnold in [4] obtained asymptotic expansion for a bounded extended convex set B:

$$\Pr(\boldsymbol{X} \in B) = J_1 + J_2 + O(n^{-1}),$$

where

$$J_1 = J_1(B) = \iint_B \phi(\boldsymbol{x})\left\{1 + \frac{1}{\sqrt{n}}\,h_1(\boldsymbol{x}) + \frac{1}{n}\,h_2(\boldsymbol{x})\right\}\,d\boldsymbol{x},$$

with

$$
\begin{aligned}
h_1(x) &= -\frac{1}{2}\sum_{j=1}^{3}\frac{x_j}{p_j} + \frac{1}{6}\sum_{j=1}^{3} x_j\left(\frac{x_j}{p_j}\right)^2, \\
h_2(x) &= \frac{1}{2}\,h_1(x)^2 + \frac{1}{12}\left(1 - \sum_{j=1}^{3}\frac{1}{p_j}\right) \\
&\quad + \frac{1}{4}\sum_{j=1}^{3}\left(\frac{x_j}{p_j}\right)^2 - \frac{1}{12}\sum_{j=1}^{3} x_j\left(\frac{x_j}{p_j}\right)^3;
\end{aligned}
$$

and

$$
\begin{aligned}
J_2 = J_2(B) = &-\frac{1}{n}\sum_{y\in L_2}\chi_{B_1}(y)\left[S_1\left(\sqrt{n}x + p_1 n\right)\phi(x,y)\right]_{\lambda_1(y)}^{\theta_1(y)} \\
&-\frac{1}{\sqrt{n}}\int_{-\infty}^{\infty}\chi_{B_2}(x)\left[S_1\left(\sqrt{n}y + p_2 n\right)\phi(x,y)\right]_{\lambda_2(x)}^{\theta_2(x)} dx, \qquad (3)
\end{aligned}
$$

with

$$
L_2 = \{y : y = \frac{1}{\sqrt{n}}(m - np_2), m \in \mathbb{Z}\}, \qquad (4)
$$

$$
S_1(x) = x - [x] - \frac{1}{2} \quad \text{and} \quad [h(x)]_{\lambda(y)}^{\theta(y)} = h(\theta(y)) - h(\lambda(y)),
$$

here $\chi_A(x)$ is an indicator function of A; a function $\phi(x,y)$ is a probability density function of standard normal distribution in $\mathbb{R}^2$ and θ_1, λ_1, θ_2, λ_2 are continuous functions from definition 1 for the set B.

M. Siotani and Y. Fujikoshi in [3] showed, that for $\lambda = 0$ and $\lambda = -1/2$ one has:

$$
J_1(B^\lambda) = K_2(c) + O(n^{-1}), \qquad (5)
$$

$$
J_2(B^\lambda) = (N^\lambda - nV^\lambda)\, e^{-\frac{c}{2}} / (2\pi n)\sqrt{\prod\nolimits_{j=1}^{3} p_j} + o(1), \qquad (6)
$$

$$
V^\lambda = V^1 + O\left(\frac{1}{n}\right),
$$

where $K_2(c)$ is the distribution function of chi-square distribution with two degrees of freedom, N^λ is a number of points from the lattice L lying in B^λ,

V^λ is an area of B^λ. These results were extended by T. Read to the case of arbitrary $\lambda \in \mathbb{R}$. It follows from theorem 3.1 in [2] that

$$\Pr\left(T_\lambda < c\right) = \Pr\left(\chi_2^2 < c\right) + J_2(B^\lambda) + O\left(n^{-1}\right),$$

and for $J_2(B^\lambda)$ the representation (6) holds. Thus, the initial problem to find rate of convergence for approximation of $\Pr\left(T_\lambda < c\right)$ is reduced to the problem of finding order of $J_2(B^\lambda)$.

Since B^λ is an extended convex set (it will be shown in lemmas 5 and 8), we can apply Yarnold's result (see [4], p. 1571) for $J_2(B^\lambda)$ and get:

$$J_2(B^\lambda) = O\left(n^{-1/2}\right).$$

In the present paper we prove a better estimate

Theorem 1. *For all $\lambda \in \mathbb{R}$ we have*

$$J_2(B^\lambda) = O\left(n^{-100/146}(\log n)^{315/146}\right). \tag{7}$$

The proof is divided into two main parts. In the first part (see Section 2) we estimate the order of approximation of $J_2(B^\lambda)$ by first summand in (6). In the second part (see Sections $3, 4$ and 5) we show that Huxley results can be applied to the set B^λ, and therefore, finally we get the order of $J_2(B^\lambda)$.

2 Expression for $J_2(B^\lambda)$

Let $\hat{\theta}_1$ and $\hat{\lambda}_1$ be the functions from definition 1 for ellipse
$B^1 = \{(x,y) : T_1(x,y) < c\}$ with

$$T_1(x,y) = \left(\frac{1}{p_1} + \frac{1}{p_3}\right)x^2 + \frac{2}{p_3}xy + \left(\frac{1}{p_2} + \frac{1}{p_3}\right)y^2$$

and let B_1^1 be domain of definition of the functions.

Lemma 1. *Lebesgue measure of a set $B_1^\lambda \setminus B_1^1$ is of order $O\left(n^{-1/2}\right)$.*

Proof. Solving an equation

$$T_1(x,y) = c$$

with respect to x, we find precise expressions for $\hat{\theta}_1$ and $\hat{\lambda}_1$:

$$\hat{\theta}_1(y) = -\frac{p_1 y}{p_1 + p_3} + \frac{\sqrt{p_1 p_2 p_3}\sqrt{-y^2 + c p_2(p_1 + p_3)}}{p_2(p_1 + p_3)},$$
$$\hat{\lambda}_1(y) = -\frac{p_1 y}{p_1 + p_3} - \frac{\sqrt{p_1 p_2 p_3}\sqrt{-y^2 + c p_2(p_1 + p_3)}}{p_2(p_1 + p_3)}.$$

Therefore the domain of definition of these functions can be written as :

$$B_1^1 = \left[-\sqrt{c p_2(p_1 + p_3)}, \sqrt{c p_2(p_1 + p_3)}\right]. \tag{8}$$

In lemmas 5 and 8 below we show that B^λ is a convex set with a smooth boundary. Therefore, there exist points on Y-axis such that the straight lines passing through the points and parallel to X-axis are the tangent lines to the curve defined by $T_\lambda(x, y) = c$. These points have the minimal $y_{\min}$ and maximal $y_{\max}$ values of the second component among all points of the curve. Thus, these extremal points are the left and right points correspondingly of an interval B_1^λ. Since for any $x, y \in B^\lambda$ starting from some $n = n(y)$ we have

$$\frac{\partial^2 T_\lambda}{\partial x^2}(x, y) > 0,$$

the function $T_\lambda(x, y)$ reaches its minimum at the point of tangency when

$$\frac{\partial T_\lambda}{\partial x}(x, y) = 0.$$

Solving the equation with respect to y, we get that the points of the curve $T_\lambda(x, y) = c$ with second components $y_{\min}$ and $y_{\max}$ lie on the straight line

$$x = -\frac{p_1 y}{p_1 + p_3}.$$

Substituting this expression into equation $T_\lambda(x, y) = c$, and expanding the left-hand side by Taylor formula we obtain

$$y_{\min} = -\sqrt{c p_2(p_1 + p_3)} + O\left(n^{-1/2}\right),\ y_{\max} = \sqrt{c p_2(p_1 + p_3)} + O\left(n^{-1/2}\right).$$

Therefore the set B_1^λ has the form

$$B_1^\lambda = \left[-\sqrt{c p_2(p_1 + p_3)} + O\left(n^{-1/2}\right), \sqrt{c p_2(p_1 + p_3)} + O\left(n^{-1/2}\right)\right]. \tag{9}$$

Now lemma follows from (8) and (9). □

Put

$$B_{1-}^1 = \left[-\sqrt{cp_2(p_1+p_3)} + n^{-1/2}, \sqrt{cp_2(p_1+p_3)} - n^{-1/2}\right]. \tag{10}$$

R e m a r k 3. Exactly two points of the lattice L_2 lie in a set $B_1^1 \setminus B_{1-}^1$ (see (4), (8) and (10)).

R e m a r k 4. Lebesgue measure of the set $B_1^\lambda \setminus B_{1-}^1$ is of order $O\left(n^{-1/2}\right)$ (see lemma 1, (8) and (10)).

R e m a r k 5. The set $B_1^\lambda \setminus B_{1-}^1$ is a union of no more than two semi-intervals.

Lemma 2. *Let θ_1 and λ_1 be the functions from definition 1 for the set B^λ (see (1)). There exist constants $c_1 > 0$ and $c_2 > 0$ such that θ_1 and λ_1 satisfy the following inequalities*

$$\left|\theta_1(y) - \hat{\theta}_1(y)\right| \le c_1 n^{-1/4}, \quad \left|\lambda_1(y) - \hat{\lambda}_1(y)\right| \le c_2 n^{-1/4} \tag{11}$$

for all $y \in B_1^\lambda \cap B_{1-}^1$ and $n \ge N = \lceil (cp_2(p_1+p_3))^{-1} \rceil$.

Proof. Expanding in the equation

$$T_\lambda(\theta_1(y), y) = c$$

the left-hand side by powers of n we get

$$T_1(\theta_1(y), y) + R(y)n^{-1/2} = c, \tag{12}$$

with

$$|R(y)| \le c_3. \tag{13}$$

We can solve (12) with respect to $\theta_1(y)$ and get

$$\left|\theta_1(y) - \hat{\theta}_1(y)\right| = \frac{\sqrt{p_1p_2p_3}|R(y)|}{\sqrt{n}}$$

$$\times \left|\sqrt{-y^2 + \left(c - \frac{R(y)}{\sqrt{n}}\right) p_2(p_1+p_3)} + \sqrt{-y^2 + cp_2(p_1+p_3)}\right|^{-1}. \tag{14}$$

It follows from (10) that for all $y \in B^1_{1-}$ we have

$$y^2 \le cp_2(p_1+p_3) - \frac{2\sqrt{cp_2(p_1+p_3)}}{\sqrt{n}} + \frac{1}{n} \tag{15}$$

By (13) – (15) we obtain for all $n \ge N = [(cp_2(p_1+p_3))^{-1}]$:

$$\left|\theta_1(y) - \hat{\theta}_1(y)\right| = \frac{\sqrt{p_1 p_2 p_3} c_3}{c^{1/4}(p_2(p_1+p_3))^{1/4}} n^{-1/4}. \tag{16}$$

This implies the first inequality in (11).

We prove similarly the second inequality in (11).

Lemma is proved. □

R e m a r k 6. Similar bounds could be obtained for the functions θ_2 and λ_2.

Statement 1. *We can write $J_2(B^\lambda)$ defined by (3), in the form*

$$J_2(B^\lambda) = \frac{d}{n}(N^\lambda - nV^\lambda) + O(n^{-3/4}), \tag{17}$$

where d is a positive constant.

Proof. We consider terms in the expression (3) separately:

$$\begin{aligned} J_{2,1} &= \frac{1}{n} \sum_{y \in L_2} \chi_{B_1^\lambda}(y) \left[S_1\left(\sqrt{n}x + p_1 n\right)\phi(x,y)\right]_{\lambda_1(y)}^{\theta_1(y)}, \\ J_{2,2} &= \frac{1}{\sqrt{n}} \int_{-\infty}^{\infty} \chi_{B_2^\lambda}(x) \left[S_1\left(\sqrt{n}y + p_2 n\right)\phi(x,y)\right]_{\lambda_2(x)}^{\theta_2(x)} dx. \end{aligned} \tag{18}$$

Then

$$J_2(B^\lambda) = -(J_{2,1} + J_{2,2}). \tag{19}$$

Using identity $B_1^\lambda = (B_1^\lambda \cap B^1_{1-}) \bigcup (B_1^\lambda \setminus B^1_{1-})$, we can rewrite $J_{2,1}$ as

$$\begin{aligned} J_{2,1} = \frac{1}{n} \sum_{y \in L_2} \chi_{B_1^\lambda \cap B^1_{1-}}(y) \left[S_1\left(\sqrt{n}x + p_1 n\right)\phi(x,y)\right]_{\lambda_1(y)}^{\theta_1(y)} \\ + \frac{1}{n} \sum_{y \in L_2} \chi_{B_1^\lambda \setminus B^1_{1-}}(y) \left[S_1\left(\sqrt{n}x + p_1 n\right)\phi(x,y)\right]_{\lambda_1(y)}^{\theta_1(y)}. \end{aligned} \tag{20}$$

The lattice L_2 has a step $n^{-1/2}$. Therefore, according to Remarks 4 and 5 there are at most $O(1)$ points of the lattice in the set $B_1^\lambda \setminus B_{1-}^1$. Hence, the second summand in (20) is of order $O\left(n^{-1}\right)$. Then using Lagrange's formula we get

$$\begin{aligned} J_{2,1} = \frac{1}{n} \sum_{y \in L_2 \cap B_1^\lambda \cap B_{1-}^1} S_1\left(\sqrt{n}\theta_1(y) + p_1 n\right) \frac{\partial \phi}{\partial x}(\xi_1(y), y) \left(\theta_1(y) - \hat{\theta}_1(y)\right) \\ + \frac{1}{n} \sum_{y \in L_2 \cap B_1^\lambda \cap B_{1-}^1} S_1\left(\sqrt{n}\lambda_1(y) + p_1 n\right) \frac{\partial \phi}{\partial x}(\xi_2(y), y) \left(\hat{\lambda}_1(y) - \lambda_1(y)\right) \\ + \frac{1}{n} \sum_{y \in L_2 \cap B_1^\lambda \cap B_{1-}^1} d\left[S_1\left(\sqrt{n}x + p_1 n\right)\right]_{\lambda_1(y)}^{\theta_1(y)} \\ + \frac{1}{n} \sum_{y \in L_2 \cap (B_1^\lambda \setminus B_{1-}^1)} \left[S_1\left(\sqrt{n}x + p_1 n\right) \phi(x,y)\right]_{\lambda_1(y)}^{\theta_1(y)}, \end{aligned}$$

where $\xi_1(y)$ and $\xi_2(y)$ are some functions defined on $B_1^\lambda \cap B_{1-}^1$. Additionally let us write

$$\begin{aligned} & \sum_{y \in L_2 \cap B_1^\lambda \cap B_{1-}^1} d\left[S_1\left(\sqrt{n}x + p_1 n\right)\right]_{\lambda_1(y)}^{\theta_1(y)} \\ = & \sum_{y \in L_2 \cap B_1^\lambda} d\left[S_1\left(\sqrt{n}x + p_1 n\right)\right]_{\lambda_1(y)}^{\theta_1(y)} - \sum_{y \in L_2 \cap (B_1^\lambda \setminus B_{1-}^1)} d\left[S_1\left(\sqrt{n}x + p_1 n\right)\right]_{\lambda_1(y)}^{\theta_1(y)} \end{aligned}$$

By Remark 4, lemma 2 and boundness of the functions S_1 and ϕ we conclude that

$$J_{2,1} = \frac{1}{n} \sum_{x_2 \in L_2 \cap B_1^\lambda} d\left[S_1\left(\sqrt{n}x + p_1 n\right)\right]_{\lambda_1(y)}^{\theta_1(y)} + O\left(n^{-3/4}\right). \tag{21}$$

Applying the same arguments to (18), we can rewrite it in the form

$$J_{2,2} = \frac{1}{\sqrt{n}} \int_{B_2^\lambda} d\left[S_1\left(\sqrt{n}y + p_2 n\right)\right]_{\lambda_2(x)}^{\theta_2(x)} dx + O\left(n^{-3/4}\right). \tag{22}$$

By (19), (21) and (22) we obtain

$$\begin{aligned} -J_2(B^\lambda) = \frac{1}{n} \sum_{y \in L_2 \cap B_1^\lambda} d\left[S_1\left(\sqrt{n}x + p_1 n\right)\right]_{\lambda_1(y)}^{\theta_1(y)} \\ + \frac{1}{\sqrt{n}} \int_{B_2^\lambda} d\left[S_1\left(\sqrt{n}y + p_2 n\right)\right]_{\lambda_2(x)}^{\theta_2(x)} dx + O\left(n^{-3/4}\right). \end{aligned} \tag{23}$$

Since we have in (23) the same constant d in the sum and integral, we can apply now the Yarnold's arguments (see [4]) and get

$$J_2(B^\lambda) = \frac{d}{n}\left(N^\lambda - nV^\lambda\right) + O\left(n^{-3/4}\right).$$

The statement is proved. □

3 Convexity of the set B^λ

Definition 2. A quadratic form in variables h_1, h_2, ..., h_m:

$$\Phi(h_1, h_2, \dots, h_m) = \sum_{i=1}^{m}\sum_{k=1}^{m} a_{ik}h_ih_k \tag{24}$$

is called *positive definite*, when for all values h_1, h_2, ..., h_m, not equal to zero simultaneously, the form takes positive values only.

Definition 3. We call a matrix

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1m} \\ a_{21} & a_{22} & \dots & a_{2m} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mm} \end{pmatrix} \tag{25}$$

by *matrix of quadratic form* (24).

Theorem. *(Sylvester's theorem) In order that a quadratic form (24) with symmetric matrix (25) is positive definite it is necessary and sufficient that the main minors of the matrix (25) are positive.*

Proof. See e.g. [8], ch. XVII, §102, theorem 102.4. □

Lemma 3. *Let a function $f(x)$, defined on a convex set Q, be two times differentiable . In order that the function is strictly convex on the set Q, it is sufficient that a second differential d^2f of the function is a positive definite quadratic form in all points of Q.*

Proof. See e.g. [7], ch.14, §7, lemma 2. □

Lemma 4. *The function $T_\lambda(x,y)$, defined in (2), is strictly convex on a set*

$$Q = \{(x,y) : x > -\sqrt{n}p_1, y > -\sqrt{n}p_2, x+y < \sqrt{n}p_3\}.$$

Proof. The set Q is convex because it is just an open triangular. Let us compute partial derivatives of the second order for $T_\lambda(x,y)$:

$$\frac{\partial^2 T_\lambda}{\partial x^2} = 2\left[\frac{1}{p_1}\left(1+\frac{x}{\sqrt{n}p_1}\right)^{\lambda-1} + \frac{1}{p_3}\left(1-\frac{x+y}{\sqrt{n}p_3}\right)^{\lambda-1}\right],$$
$$\frac{\partial^2 T_\lambda}{\partial y^2} = 2\left[\frac{1}{p_2}\left(1+\frac{y}{\sqrt{n}p_2}\right)^{\lambda-1} + \frac{1}{p_3}\left(1-\frac{x+y}{\sqrt{n}p_3}\right)^{\lambda-1}\right],$$
$$\frac{\partial^2 T_\lambda}{\partial x \partial y} = \frac{2}{p_3}\left(1-\frac{x+y}{\sqrt{n}p_3}\right)^{\lambda-1} = \frac{\partial^2 (T_\lambda)}{\partial y \partial x}.$$

All computed derivatives are continuous in Q. Therefore, the function $T_\lambda(x,y)$ is two times differentiable in Q. By lemma 3 it is sufficient to show that $d^2(T_\lambda)$ is positive definite quadratic form. By Sylvester's theorem it is sufficient to show that main minors of a matrix

$$A = \begin{pmatrix} \frac{\partial^2(T_\lambda)}{\partial x^2} & \frac{\partial^2(T_\lambda)}{\partial x \partial y} \\ \frac{\partial^2(T_\lambda)}{\partial y \partial x} & \frac{\partial^2(T_\lambda)}{\partial y^2} \end{pmatrix}$$

are positive.

It is clear that for all $(x,y) \in Q$ the main first order minor $A_1 = \partial^2(T_\lambda)/\partial x^2$ is positive. The main second order minor equals

$$\begin{aligned} A_2 &= \frac{\partial^2(T_\lambda)}{\partial x^2}\frac{\partial^2(T_\lambda)}{\partial y^2} - \frac{\partial^2(T_\lambda)}{\partial x \partial y}\frac{\partial^2(T_\lambda)}{\partial y \partial x} \\ &= 4\left[\frac{(ab)^{\lambda-1}}{p_1p_2} + \frac{(ac)^{\lambda-1}}{p_1p_3} + \frac{(bc)^{\lambda-1}}{p_2p_3}\right] > 0, \end{aligned}$$

where $a = 1 + x/\sqrt{np_1} > 0$, $b = 1 + y/\sqrt{np_2} > 0$ and $c = 1 - (x+y)/\sqrt{np_3} > 0$.

Lemma is proved. □

Lemma 5. *B^λ is a strictly convex set.*

Proof. Fix any

$$\boldsymbol{x_1} = (x_1, y_1) \in B^\lambda,\ \boldsymbol{x_2} = (x_2, y_2) \in B^\lambda \text{ and } t \in [0,1].$$

Then $T_\lambda(\boldsymbol{x_1}) < c$, $T_\lambda(\boldsymbol{x_2}) < c$. It follows from lemma 4 that $T_\lambda(x,y)$ is strictly convex function on Q. Therefore,

$$\begin{aligned} T_\lambda(\boldsymbol{x_1} + t(\boldsymbol{x_2} - \boldsymbol{x_1})) &< T_\lambda(\boldsymbol{x_1}) + t(T_\lambda(\boldsymbol{x_2}) - T_\lambda(\boldsymbol{x_1})) \\ &= (1-t)T_\lambda(\boldsymbol{x_1}) + tT_\lambda(\boldsymbol{x_2}) < (1-t)c + tc = c. \end{aligned}$$

Hence, $\boldsymbol{x_1} + t(\boldsymbol{x_2} - \boldsymbol{x_1}) \in B^\lambda$, and therefore B^λ is convex set. Repeating these arguments for any pair of points from the boundary of B^λ we get that the set is strictly convex.

Lemma is proved. □

4 Smoothness of the curve $T_\lambda(x,y) = c$

Let us consider function

$$U(r,t) = T_\lambda(r\cos t, r\sin t) - c, \tag{26}$$

on a set

$$\begin{aligned} &S = (0, +\infty) \times [0, 2\pi] \\ &\cap \{(r,t) : r\cos t > -\sqrt{n}p_1, r\sin t > -\sqrt{n}p_2, r\cos t + r\sin t < \sqrt{n}p_3\}. \end{aligned} \tag{27}$$

Lemma 6. *We have*

$$\exists s, N \colon \forall (r,t) \in \partial B^\lambda,\ n \geqslant N \quad \frac{\partial U(r,t)}{\partial r} \geqslant s > 0. \tag{28}$$

Proof. We expand a partial derivative of U in powers of n:

$$\begin{aligned} \frac{\partial U(r,t)}{\partial r} &= 2r\left(\cos^2 t\left(\frac{1}{p_1} + \frac{1}{p_3}\right) + \sin^2 t\left(\frac{1}{p_2} + \frac{1}{p_3}\right) + \frac{2\cos t \sin t}{p_3}\right) \\ &\quad + O\left(\frac{1}{\sqrt{n}}\right). \end{aligned}$$

It is clear that on the boundary of the curve $U(r,t) = 0$ there exists r_1 such that for all t we have $r(t) \geqslant r_1$. Since B^λ is bounded and due to the structure of the function $U(r,t)$ infinitely differentiable on $(r,t) \in [0,r_0] \times [0,2\pi]$, the given order $O\left(1/\sqrt{n}\right)$ of remainder term is uniform with respect to t. Changing to the double trigonometric variable and then using formula of the cosine of additional variable we get a lower bound for the derivative

$$\frac{1}{2}\left(\frac{1}{p_1}+\frac{1}{p_2}+\frac{2}{p_3}\right)+\sqrt{\frac{(1/p_1-1/p_2)^2}{4}+\frac{1}{p_3^2}}\cos(2t+\phi_0)$$
$$\geq \frac{1}{2}\left(\frac{1}{p_1}+\frac{1}{p_2}+\frac{2}{p_3}\right)-\sqrt{\left(\frac{1}{2p_1}\right)^2+\left(\frac{1}{2p_2}\right)^2+\left(\frac{1}{p_3}\right)^2-\frac{1}{2p_1p_2}}$$
$$> \frac{1}{2p_1}+\frac{1}{2p_2}+\frac{1}{p_3}-\sqrt{\left(\frac{1}{2p_1}\right)^2+\left(\frac{1}{2p_2}\right)^2+\left(\frac{1}{p_3}\right)^2} > 0.$$

Lemma is proved. □

Theorem. *(Existence and differentiability of an implicit function) Let a function $F(x,y)$ be k times differentiable in some neighborhood of a point (x_0,y_0) in $\mathbb{R}^2$. Assume that a partial derivative $\partial F/\partial y$ is continuous at (x_0,y_0). If*

$$F(x_0,y_0)=0, \quad \text{and} \quad \frac{\partial F}{\partial y}(x_0,y_0)\neq 0,$$

then for any sufficiently small positive number ε there exists such neighborhood of x_0 in $\mathbb{R}$, that in this neighborhood there exists a unique function $y=\phi(x)$ satisfying $|y-y_0|<\varepsilon$ which is a solution of the equation

$$F(x,y)=0,$$

*and $\phi(x)$ is continuous and k times differentiable function in the mentioned neighborhood
of x_0.*

Proof. See e.g. [10], ch. 1, §1. □

Lemma 7. *Let (r_0,t_0) be a point in S where the function $U(r,t)$ equals 0. Then for any sufficiently small positive number ε there exists a neighborhood of t_0 such that in the neighborhood there exists a unique function $r=r(t)$ satisfying $|r-r_0|<\varepsilon$ that is a solution of the equation*

$$U(r,t)=0,$$

and $r(t)$ is a continuous and five times differentiable function in the mentioned neighborhood of t_0.

Proof. Let (r_0, t_0) be a point in S, where the function $U(r, t)$ equals 0. Since S is an open set, there exists a neighborhood of (r_0, t_0) lying completely in S. The function $U(r, t)$ is infinitely differentiable in the mentioned neighborhood. Hence, the partial derivative $\partial U/\partial r$ is continuous at (r_0, t_0). By lemma 6 the partial derivative $\partial U/\partial r$ does not equal zero at (r_0, t_0). Therefore, $U(r, t)$ satisfies all conditions of the previous theorem at the point (r_0, t_0). Thus, the lemma follows from the theorem above. $\square$

Lemma 8. *For the curve*

$$T_\lambda(x, y) = c \tag{29}$$

there exists four times differentiable parametrization in the form

$$x = x(t) = r(t)\cos t, \quad y = y(t) = r(t)\sin t$$

for $t \in [0, 2\pi]$.

Proof. By lemma 5 the set $B^\lambda = \{(x, y) : T_\lambda(x, y) < c\}$ is convex. Moreover, the origin of coordinates lies in B^λ, because $T_\lambda(0, 0) = 0 < c$. Therefore, for any $t_0 \in [0, 2\pi]$ a half-line starting from the origin under angle t_0 to X-axis intersects the curve (29) in one point (x_0, y_0) only. Let us turn to the polar system of coordinates:

$$x = r\cos t, \qquad y = r\sin t.$$

Then the point (x_0, y_0) turns into (r_0, t_0) where $r_0 = \sqrt{x_0^2 + y_0^2}$. Since (x_0, y_0) lies on the curve (29), we have

$$U(r_0, t_0) = T_\lambda(r_0 \cos t_0, r_0 \sin t_0) - c = T_\lambda(x_0, y_0) - c = 0.$$

Therefore, by lemma 7 in some neighborhood of t_0 there exists a unique function $r = r(t)$ as the solution of $U(r, t) = 0$. Moreover, $r(t)$ is continuous and five times differentiable in this neighborhood. Let

$$x(t) = r(t)\cos t, \qquad y(t) = r(t)\sin t.$$

Then in the indicated neighborhood of t_0 we have

$$T_\lambda(x(t), y(t)) = T_\lambda(r(t)\cos t, r(t)\sin t) = U(r(t), t) + c = c,$$

and $x(t), y(t)$ are continuous and five times differentiable functions in this neighborhood. Therefore, they are four times continuously differentiable in the neighborhood, and hence they give the desired parametrization of the curve (29) in the neighborhood of t_0.

Since we choose t_0 arbitrarily, the desired parametrization exists on the whole interval $[0, 2\pi]$.

Lemma is proved. □

Corollary 1. *Radius of curvature of the curve (29) is non-zero on the entire curve.*

Proof. Let $x(t), y(t)$ be parametrization of the curve (29) from lemma 8. We show that

$$(x'(t))^2 + (y'(t))^2 \neq 0 \qquad \text{for all } \ t \in [0, 2\pi]. \tag{30}$$

In fact, assume that there exists $t_0 \in [0, 2\pi]$ such that $(x'(t_0))^2 + (y'(t_0))^2 = 0$. Then

$$r'^2(t_0) + r^2(t_0) = 0.$$

Therefore,

$$r(t_0) = 0 \ \Rightarrow \ \begin{cases} x(t_0) = 0, \\ y(t_0) = 0. \end{cases} \ \Rightarrow \ T_\lambda(x(t_0), y(t_0)) = 0,$$

which contradicts the fact that $x(t), y(t)$ is a parametrization of the curve (29).

Furthermore, according to the formula for radius of curvature we have

$$\rho = \frac{((x')^2 + (y')^2)^{3/2}}{x'y'' - y'x''}, \tag{31}$$

which, together with (30), implies the statement of this corollary.

□

Definition 4. A curve $\{x(t), y(t)\}$, $t \in [a, b]$ is called *smooth*, when the functions $x(t), y(t)$ are smooth on $[a, b]$.

Definition 5. A smooth curve $\{x(t), y(t)\}$, $t \in [a, b]$ is called *regular*, when vector $(x'(t), y'(t))^T$ does not equal zero everywhere on $[a, b]$.

Definition 6. A parameter l of a curve $\{x(l), y(l)\}$ is called *natural*, if the length of the curve equals $(b_1 - a_1)$ as l runs from a_1 to $b_1 > a_1$.

Lemma 9. *1) If $l \in [a,b]$ on the curve $\{x(l), y(l)\}$ is a natural parameter, then*

$$\sqrt{(x'(l))^2 + (y'(l))^2} = 1$$

at all points where continuous derivatives $x'(l), y'(l)$ exist.

2) For any regular curve there exists a natural parameter.

Proof. See e.g. [9], ch. 1, §1, lemma 2. □

Corollary 2. *Radius of curvature of the curve (29) is continuous on the curve.*

Proof. Let $x(t), y(t)$ be the parametrization of the curve (29) from lemma 8. Now we show that

$$x'y'' - y'x'' \neq 0 \qquad \text{for all } t \in [0, 2\pi]. \tag{32}$$

At first, we prove that $(x'')^2 + (y'')^2 \neq 0$ everywhere on $[0, 2\pi]$. Assume that just the opposite is true, that is for some $t_0 \in [0, 2\pi]$ we have:

$$(x''(t_0))^2 + (y''(t_0))^2 = 0.$$

Then using expressions for $x(t)$ and $y(t)$ from lemma 8 we get:

$$4(r'(t_0))^2 + (r''(t_0) - r(t_0))^2 = 0,$$

and hence,

$$\begin{cases} r'(t_0) = 0, \\ r''(t_0) = r(t_0). \end{cases} \tag{33}$$

Furthermore, by differentiating twice the identity

$$U(r(t), t) = 0$$

at point t_0 and taking into account (33) we get

$$\frac{2r^2(t_0)\sin^2 t_0}{p_1\left(1 + r(t_0)\cos t_0/\sqrt{n}p_1\right)^{1-\lambda}} + \frac{2r^2(t_0)\cos^2 t_0}{p_2\left(1 + r(t_0)\sin t_0/\sqrt{n}p_2\right)^{1-\lambda}}$$
$$+ \frac{2(-r(t_0)\sin t_0 + r(t_0)\cos t_0)^2}{p_3\left(1 - (r(t_0)\cos t_0 + r(t_0)\sin t_0)/\sqrt{n}p_3\right)^{1-\lambda}} = 0. \tag{34}$$

Here the denominators of each fraction are positive due to the domain of definition for $U(r,t)$ (see (27)). Therefore, each of the summands in (34) is equal to zero. Consequently,

$$\cos t_0 = \sin t_0 = 0,$$

but this contradicts the Pythagorean trigonometric identity. Thus,

$$(x'')^2 + (y'')^2 \neq 0$$

everywhere on the curve.

From lemma 8 and (30) we conclude that the curve (29) is regular and due to lemma 9 allows natural parametrization of the form $x = \chi(l)$, $y = \gamma(l)$. It can be shown that in this case the vectors $(\chi', \gamma')^T$, $(\chi'', \gamma'')^T$ are also nonzero everywhere on $l \in [0, L]$ where L is the length of the curve (29) (it can be easily shown by the rule of contraries using the fact that the mapping $l :\ [0, 2\pi] \to [0, L]$ defined by the formula

$$l(t) = \int_0^t \sqrt{x'^2(\tau) + y'^2(\tau)} d\tau \tag{35}$$

is smooth and invertible). But then lemma 9 implies

$$\chi'^2(l) + \gamma'^2(l) = 1.$$

Differentiating this identity with respect to l we obtain:

$$\chi'(l)\chi''(l) + \gamma'(l)\gamma''(l) = 0,$$

and, consequently, the vectors $(\chi', \gamma')^T$ and (χ'', γ'') are orthogonal. Therefore, the determinant

$$\begin{vmatrix} \chi'(l) & \chi''(l) \\ \gamma'(l) & \gamma''(l) \end{vmatrix} \neq 0 \ \Leftrightarrow\ \chi'(l)\gamma''(l) - \gamma'(l)\chi''(l) \neq 0. \tag{36}$$

Thus, since $l(t)$ defined in (35) is one-to-one mapping, (32) holds. Hence, using the formula for the radius of curvature (31) we obtain the statement of the corollary.

□

Corollary 3. *The radius of curvature on the curve (29) is twice continuously differentiable with respect to the tangent angle everywhere on that curve.*

Proof. Let $\chi = \chi(l)$, $\gamma = \gamma(l)$ be a natural parametrization of the curve (29). Then it follows from lemma 8 and from the smoothness and invertibility of the mapping (35) that $\chi(l)$ and $\gamma(l)$ are four times continuously differentiable functions. Further, let ρ be the radius of curvature of the curve (29) and ψ be a tangent angle. Then

$$\frac{d\rho}{d\psi} = \frac{d\rho}{dl}\frac{dl}{d\psi} = \rho\frac{d\rho}{dl} = \frac{1}{2}\frac{d\rho^2}{dl} = \frac{1}{2}\frac{d\left(\frac{\left(\chi'^2+\gamma'^2\right)^3}{(\chi'\gamma''-\gamma'\chi'')^2}\right)}{dl}$$
$$= \frac{3}{2}\frac{\left(\chi'^2+\gamma'^2\right)^2(2\chi'\gamma''+2\gamma'\chi'')}{(\chi'\gamma''-\gamma'\chi'')^2} - \frac{(\chi'^2+\gamma'^2)^3(\chi'\gamma'''-\gamma'\chi''')}{(\chi'\gamma''-\gamma'\chi'')^3}. \quad (37)$$

Due to the smoothness of the functions $\chi(l)$ and $\gamma(l)$ and property (36) we conclude that the radius of curvature ρ is continuously differentiable everywhere on the curve (29).

Similarly,

$$\frac{d^2\rho}{d\psi^2} = \frac{d}{d\psi}\left(\frac{d\rho}{d\psi}\right) = \frac{1}{2}\rho\frac{d\left(\frac{d\rho^2}{dl}\right)}{dl} \quad (38)$$

Without giving the exact formula for the second derivative with respect to the tangent angle it can be easily seen that the derivative is continuous due to the constraints imposed on $\chi(l), \gamma(l)$ and the fact that in the denominator of the resultant expression we will again get $\chi'\gamma''-\gamma'\chi''$ raised to some power.

Corollary is proved. □

5 Applying Huxley's theorem to the set B^λ

Theorem 2. *(Huxley, 1993) Let B be a Euclidean plane domain of area A, bounded by a simple closed curve C, composed of finitely many pieces C_i, which are three times continuously differentiable in the following sense. The radius of curvature ρ is continuous and non-zero on each piece C_i, and ρ is continuously differentiable with respect to the tangent angle ψ. Let MB denote the set formed by expanding B linearly by a factor M. Then for any isometric embedding of MB in the Euclidean plane the number of integer points (m, n) in MB is*

$$AM^2 + O\left(IM^{46/73}(\log M)^{315/146}\right), \quad (39)$$

where I is a number depending on the curve C, but not on M or on the embedding of MB.

If in addition the pieces C_i are four times differentiable, in the sense that ρ is twice continuously differentiable with respect to tangent angle ψ, then we may take

$$\begin{aligned} I &= \sum_i \min_{C_i} \left(1 + \frac{1}{\rho^2}\left(\frac{d\rho}{d\psi}\right)^2\right)^{-69/146} \rho^{46/73} \\ &+ \sum_i \int_{C_i} \left(1 + \frac{|\rho\, d^2\rho/d\psi^2|}{\rho^2 + (d\rho/d\psi)^2}\right) \\ &\times \left(1 + \frac{1}{\rho^2}\left(\frac{d\rho}{d\psi}^2\right)^2\right)^{-69/146} \left|\frac{d\rho}{d\psi}\right| \rho^{-33/73} d\psi, \end{aligned} \tag{40}$$

provided that M is so large that the bounds

$$M \geqslant \frac{1}{\rho} \quad \textit{and} \quad \frac{1}{\rho^{64}}\left|\frac{d\rho}{d\psi}\right|^{53} \leqslant M^{11}(\log M)^{387/8}$$

hold piecewise on each curve C_i.

Proof. See [5, theorems 5 and 6, pp. 294–295]. □

Now we prove lemma which shows that in our case I from theorem 2 is bounded from above by some constant not depending on n. It is necessary to note that in 2003 Huxley slightly improved the result of theorem 2. However the form of I in the improved result is such that it cannot be applied in our case.

Lemma 10. *For a sufficiently large n the radius of curvature ρ of the boundary ∂B^λ is bounded from above and separated from zero uniformly with respect to n; its first and second derivatives with respect to the tangent angle ψ are uniformly bounded from above.*

Proof. We recall that the radius of curvature and its derivatives are given by formulae (31), (37), and (38). We use the parametrization in polar coordinates from lemma 8. In this case

$$\rho = \frac{(r^2(t) + r'^2(t))^{3/2}}{|2(r'(t))^2 + r^2(t) - r'(t)r''(t)|}, \tag{41}$$

whereas the derivatives with respect to the tangent angle are expressed analogously, and expression

$$2(r'(t))^2 + r^2(t) - r'(t)r''(t) \tag{42}$$

will appear in the denominator.

Let us denote $r_n(t)$ the polar radius on ∂B^λ and $r(t)$ the polar radius on ∂B^1; the values $r'(t), r''(t), r_n'(t), r_n''(t)$ being similarly defined. Note that the exact expression of (42) for the limiting set is separated from 0. In fact, this is an ellipse rotated around the origin with the axes $a(\bar{p}, c), b(\bar{p}, c)$. For the simplest ellipse of the form

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1,$$

we substitute our parametrization and obtain

$$r(t) = \left(\frac{\cos^2 t}{a^2} + \frac{\sin^2 t}{b^2}\right)^{-1/2} = \left(\frac{1}{2}\left(\frac{1}{a^2} + \frac{1}{b^2}\right) + \frac{\cos 2t}{2}\left(\frac{1}{a^2} - \frac{1}{b^2}\right)\right)^{-1/2}, \tag{43}$$

$$r'(t) = \frac{\sin 2t}{2}\left(\frac{1}{a^2} - \frac{1}{b^2}\right)\left(\frac{1}{2}\left(\frac{1}{a^2} + \frac{1}{b^2}\right) + \frac{\cos 2t}{2}\left(\frac{1}{a^2} - \frac{1}{b^2}\right)\right)^{-3/2}, \tag{44}$$

$$\begin{aligned} r''(t) &= \Bigg(\left(\frac{1}{a^2} - \frac{1}{b^2}\right)\cos 2t \cdot \left(\frac{1}{2}\left(\frac{1}{a^2} + \frac{1}{b^2}\right) + \frac{\cos 2t}{2}\left(\frac{1}{a^2} - \frac{1}{b^2}\right)\right) \\ &\quad + \frac{3}{4}\left(\frac{1}{a^2} - \frac{1}{b^2}\right)^2 \sin^2 2t\Bigg) \\ &\quad \times\left(\frac{1}{2}\left(\frac{1}{a^2} + \frac{1}{b^2}\right) + \frac{\cos 2t}{2}\left(\frac{1}{a^2} - \frac{1}{b^2}\right)\right)^{-5/2} \\ &= \left(\frac{1}{a^2} - \frac{1}{b^2}\right)^2 \cdot \left(\frac{3}{4} - \frac{\cos^2 2t}{4} + \frac{b^2 + a^2}{b^2 - a^2}\frac{\cos 2t}{2}\right) \\ &\quad \times\left(\frac{1}{2}\left(\frac{1}{a^2} + \frac{1}{b^2}\right) + \frac{\cos 2t}{2}\left(\frac{1}{a^2} - \frac{1}{b^2}\right)\right)^{-5/2}. \end{aligned} \tag{45}$$

We see that $r(t)$ is bounded:

$$\sqrt{2}\left(\frac{1}{a^2} + \frac{1}{b^2} - \left|\frac{1}{a^2} - \frac{1}{b^2}\right|\right)^{-1/2} \geqslant r(t) \geqslant \sqrt{2}\left(\frac{1}{a^2} + \frac{1}{b^2} + \left|\frac{1}{a^2} - \frac{1}{b^2}\right|\right)^{-1/2}. \tag{46}$$

Now for (42) we have

$$\underbrace{\left(\frac{1}{2}\left(\frac{1}{a^2}+\frac{1}{b^2}\right)+\frac{\cos 2t}{2}\left(\frac{1}{a^2}-\frac{1}{b^2}\right)\right)}_{A}^{-3}\left[\frac{\sin^2 2t}{2}\left(\frac{1}{a^2}-\frac{1}{b^2}\right)^2\right.$$
$$+\left(\frac{1}{2}\left(\frac{1}{a^2}-\frac{1}{b^2}\right)+\frac{\cos 2t}{2}\left(\frac{1}{a^2}-\frac{1}{b^2}\right)\right)^2-\frac{1}{2}\left(\frac{1}{a^2}-\frac{1}{b^2}\right)^2\cdot\left(\frac{3}{2}-\frac{\cos^2 2t}{2}\right.$$
$$\left.\left.+\frac{b^2+a^2}{b^2-a^2}\cos 2t\right)\right]=A^{-3}\left(\frac{1}{4}\left(\frac{1}{a^2}+\frac{1}{b^2}\right)^2-\frac{1}{4}\left(\frac{1}{a^2}-\frac{1}{b^2}\right)^2\right)=\frac{1}{a^2b^2A^3}>0.$$

Since in polar coordinates the rotation is reduced to the transformation $t := t + c$, and the upper estimate can be made independent from t, we have proved that expression (42) for B^1 is separated from zero. It is natural to anticipate that the prelimiting set B^λ possesses the same property, at least starting from some number N, uniformly in t.

In the appendix (lemma 11) we prove the uniform convergence

$$r_n(t) \xrightarrow[n\to\infty]{} r(t).$$

We know that the derivatives of solutions $r_n(t)$, $r(t)$ are expressed through the derivatives of an implicit function with respect to its arguments t and $r(t)$. Moreover, the denominator will contain the first derivative with respect to r of the functions $T_\lambda(r,t)$ and $T_1(r,t)$ raised to some power. For instance,

$$r_n'(t) = -\frac{\partial T_\lambda(r_n(t),t)}{\partial t}\Bigg/\frac{\partial T_\lambda(r_n(t),t)}{\partial r}, \qquad r'(t) = -\frac{\partial T_1(r(t),t)}{\partial t}\Bigg/\frac{\partial T_1(r(t),t)}{\partial r}$$

From lemma 6

$$\exists N\colon \forall n \geqslant N\ \frac{\partial T_1(r(t),t)}{\partial r} \geqslant s > 0,\ \frac{\partial T_\lambda(r_n(t),t)}{\partial r} \geqslant s > 0.$$

Moreover, in lemma 6 we essentially proved the following uniform estimate

$$\frac{\partial T_\lambda(r(t),t)}{\partial r} = \frac{\partial T_1(r(t),t)}{\partial r} + O\left(\frac{1}{\sqrt{n}}\right).$$

With similar reasoning we can obtain the same result for the derivatives with respect to t:

$$\frac{\partial T_\lambda(r(t),t)}{\partial t} = \frac{\partial T_1(r(t),t)}{\partial t} + O\left(\frac{1}{\sqrt{n}}\right).$$

Therefore, it is easy to see that

$$\frac{\partial T_\lambda(r(t),t)}{\partial t}\Big/\frac{\partial T_\lambda(r(t),t)}{\partial r}=\frac{\partial T_1(r(t),t)}{\partial t}\Big/\frac{\partial T_1(r(t),t)}{\partial r}+O\left(\frac{1}{\sqrt{n}}\right). \quad (47)$$
$$\frac{\partial T_\lambda(r_n(t),t)}{\partial t}\Big/\frac{\partial T_\lambda(r_n(t),t)}{\partial r}=\frac{\partial T_1(r_n(t),t)}{\partial t}\Big/\frac{\partial T_1(r_n(t),t)}{\partial r}+O\left(\frac{1}{\sqrt{n}}\right).$$

Let us expand the difference $r_n'(t)-r'(t)$:

$$\begin{aligned}&\frac{\partial T_\lambda(r_n(t),t)}{\partial t}\Big/\frac{\partial T_\lambda(r_n(t),t)}{\partial r}-\frac{\partial T_1(r(t),t)}{\partial t}\Big/\frac{\partial T_1(r(t),t)}{\partial r}\\ &=\left(\frac{\partial T_\lambda(r_n(t),t)}{\partial t}\Big/\frac{\partial T_\lambda(r_n(t),t)}{\partial r}-\frac{\partial T_\lambda(r(t),t)}{\partial t}\Big/\frac{\partial T_\lambda(r(t),t)}{\partial r}\right)\\ &\quad+\left(\frac{\partial T_\lambda(r(t),t)}{\partial t}\Big/\frac{\partial T_\lambda(r(t),t)}{\partial r}-\frac{\partial T_1(r(t),t)}{\partial t}\Big/\frac{\partial T_1(r(t),t)}{\partial r}\right).\end{aligned}$$

Since the fraction $\frac{\partial T_1(r,t)}{\partial t}/\frac{\partial T_1(r,t)}{\partial r}$ is a smooth function independent from n, with non-zero denominator, and since variables (r,t) change in a bounded domain, we can apply (47) to get by Lagrange's theorem the following inequality

$$|r_n'(t)-r'(t)|\leqslant M\cdot|r_n(t)-r(t)|+O\left(1/\sqrt{n}\right).$$

This implies the uniform convergence of the first derivatives of polar radius. Similar arguments show the uniform convergence of the derivatives of higher order.

It follows from formulae (43),(44), (45), and (46) that the derivatives of the polar radius on ∂B^1 are bounded from above, and that the polar radius itself is bounded from both sides. Moreover, the term (42) is separate from 0. It is clear from the asymptotic properties of $r_n(t)$ and its derivatives that *the same statements* are valid for the polar radius $r_n(t)$ (together with its derivatives) of ∂B^λ, at least starting from sufficiently large N, uniformly in t. Now the statement of the lemma follows from the above arguments and formulae (41), (37), and (38). □

Corollary 4. *For sufficiently large n the set B^λ satisfies the conditions of theorem 2 with $M=\sqrt{n}$.*

6 Proof of the main result

We recall that N^λ is a number of points from the lattice L in the set B^λ. Since the lattice has $1/\sqrt{n}$ as a step, we can regard N^λ as a number of integer points in the set $\sqrt{n}B^\lambda$, which is a linear expansion of the set B^λ with the coefficient $\sqrt{n}$. Because of Corollary 4 we can apply theorem 2 to the set B^λ with the linear factor $\sqrt{n}$.

Note that in our case I from theorem 2 depends on n. However, it is bounded. This fact follows from the upper bound

$$I(n) \leqslant \min_C \rho^{46/73} + \int_C \frac{1 + \left|\frac{d^2\rho}{d\psi^2}/\rho\right|}{\rho^{33/73}} \left|\frac{d\rho}{d\psi}\right| d\psi$$

and lemma 10. Consequently, we can disregard this constant in the calculation of the error order and get from theorem 2

$$N^\lambda - nV^\lambda = O\left(n^{46/146}(\log n)^{315/146}\right). \tag{48}$$

It remains to substitute (48) into (17), and we obtain (7).

This proves theorem 1.

R e m a r k 7. We proved the uniform convergence of the polar radius $r_n(t)$ and its derivatives to their limits. We also proved that $r_n(t)$ are separated from zero uniformly with respect to n. Hence, the expressions under the signs of integration and min in (40) converge uniformly. Therefore, by Lebesgue theorem not only is $I(n)$ bounded, but it also converges to I_{B^1}.

A Proof of the uniform convergence of polar radii

Lemma 11. *Let $r_n(t)$ and $r(t)$ be the polar radii of the sets B^λ and B^1 correspondingly. Then we have*

$$|r_n(t) - r(t)| \leqslant \frac{C}{\sqrt{n}}.$$

Proof. We have

$$\begin{aligned} T_1(r_n(t),t) - T_1(r(t),t) \leqslant |T_1(r_n(t),t) - T_\lambda(r_n(t),t)| \\ + |T_\lambda(r_n(t),t) - T_\lambda(r(t),t)| + |T_\lambda(r(t),t) - T_1(r(t),t)|. \end{aligned}$$

It follows from Taylor's formula that $T_\lambda(r,t) = T_1(r,t) + O\left(1/\sqrt{n}\right)$, and the error is uniform in n due to the boundedness of the domain of definition. Therefore,

$$|T_1(r_n(t),t) - T_\lambda(r_n(t),t)| = O\left(\frac{1}{\sqrt{n}}\right),\ |T_\lambda(r(t),t) - T_1(r(t),t)| = O\left(\frac{1}{\sqrt{n}}\right).$$

Moreover, $T_\lambda(r_n(t),t) = c = T_1(r(t),t)$, and the second summand can be expressed in the form

$$|T_\lambda(r(t),t) - T_1(r(t),t)| = O\left(\frac{1}{\sqrt{n}}\right).$$

On the other hand,

$$\begin{aligned} T_1(r_n(t),t) - T_1(r(t),t) &= \frac{(r_n(t)\cos t)^2}{p_1} + \frac{(r_n(t)\sin t)^2}{p_2} + \frac{(r_n(t)(\cos t + \sin t))^2}{p_3} \\ &- \left[\frac{(r(t)\cos t)^2}{p_1} + \frac{(r(t)\sin t)^2}{p_2} + \frac{(r(t)(\cos t + \sin t))^2}{p_3}\right] \\ &= \left[\cos^2 t\left(\frac{1}{p_1} + \frac{1}{p_3}\right) + \sin^2 t\left(\frac{1}{p_2} + \frac{1}{p_3}\right) + \frac{\sin 2t}{p_3}\right](r_n^2(t) - r^2(t)). \end{aligned}$$

From lemma 6, we know that the first multiplier is uniformly separated from 0 (let us denote this multiplier by E and the corresponding lower bound by E_0). Hence, since there is a lower bound for $r(t)$, we have

$$|r_n(t) - r(t)| = O\left(\frac{1}{E(r_n(t) + r(t))\sqrt{n})}\right) = O\left(\frac{1}{E_0 r(t)\sqrt{n}}\right) = O\left(\frac{1}{\sqrt{n}}\right).$$

Lemma is proved. □

References

[1] N. A. C. Cressie, T. R. C. Read (1984). Multinomial goodness-of-fit tests, Journal of the Royal Statistical Society, Series B, Vol 46, No. 3 (1984), pp. 440-464.

[2] T. R. C. Read (1984). Closer asymptotic approximations for the distributions of the power divergence goodness-of-fit statistics., The Annals of Mathematical Statistics, 36, Part A, p. 59-69.

[3] M. Siotani and Y. Fujikoshi, Asymptotic approximations for the distributions of multinomial goodness-of-fit statistics, Hiroshima Math. J., **14** (1984), 115–124; Technical report of the Hiroshima statistical research group (1980).

[4] J. K. Yarnold, Asymptotic approximations for the probability that a sum of lattice random vectors lies in a convex set, The Annals of Mathematical Statistics 1972, Vol. 43, No. 5, 1566–1580.

[5] M. N. Huxley, Exponential sums and lattice points II, Proceedings of London Mathematical Society (3) 66 (1993) 279-301.

[6] M. N. Huxley, Exponential sums and lattice points III, Proceedings of London Mathematical Society (3) 87 (2003) 591-609.

[7] V. A. Ilyin, E. G. Pozdnyak. Foundations of Mathematical Analysis (in Russian), Part I. Moscow: FIZMATLIT, 2002.

[8] V. A. Ilyin, G. D. Kim. Linear algebra and analytical geometry. (in Russian) Moscow: Moscow State University, 1998.

[9] Taymanov I. A. Lectures on differential geometry. (in Russian) – M.-Izhevsk: Reseach centre "Regular and chaotic dynamics"; The Institute of Computer Science Research, 2006.

[10] M. M. Vainberg, V. A. Trenogin. Bifurcation theory of nonlinear equations. (in Russian) Moscow: Nauka, 1969.