UNIFORMITY IN CUBE-COVERING SYSTEMS

J. BECK, W.W.L. CHEN, AND Y. YANG

Abstract. We establish various analogs of the Kronecker–Weyl equidistribution theorem that can be considered higher-dimensional versions of results established in our earlier investigation in [1] of the discrete 2-circle problem studied in 1969 by Veech [6]. Whereas the Veech problem can be viewed as one of geodesic flow on a 2-dimensional flat surface, here we study geodesic flow in higher-dimensional flat manifolds. This is more challenging, as the overwhelming majority of the available proof techniques for non-integrable flat systems are based on arguments in dimension 2. For higher dimensions, we need a new approach.

1. Analog of the Veech 2-circle problem

We extend the idea of Veech [6] and our earlier work [1] to higher dimension. First we 2-color the unit torus $[0,1)^2$ red and green in such a way that each of the red and green parts is the union of finitely many polygons. Figure 1.1 shows two examples, where the shaded part represents red and the white part represents green. In particular, we assume that the green part has positive area.

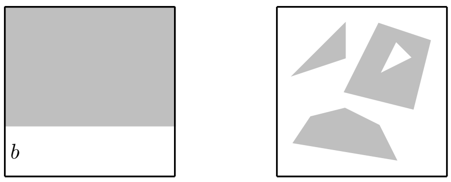

Figure 1.1: examples of 2-colorings of the torus $[0,1)^2$

Observe that in the picture on the right, one of the red (shaded) parts does not look like a polygon, but it is the union of finitely many polygons. A similar remark applies to the green (white) part.

Next, we consider a 2-torus system as shown in Figure 1.2, where each square represents the unit torus $[0,1)^2$, with identical 2-coloring.

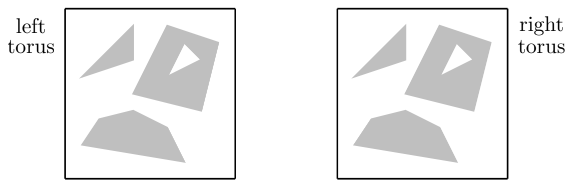

Figure 1.2: a 2-torus system with identical coloring

Let $\mathbf{v} = (1, \alpha_1, \alpha_2)$ be a Kronecker vector, and let $\mathbf{v}_0 = (\alpha_1, \alpha_2)$.

Let $\mathbf{s}_0 \in [0,1)^2$ be an arbitrary starting point, and consider the $\mathbf{v}_0$-shift sequence

$$\mathbf{s}_n = \mathbf{s}_0 + n\mathbf{v}_0, \quad n = 0, 1, 2, 3, \ldots,$$

in the unit torus $[0,1)^2$; in other words, *modulo one*. Assume that the point $\mathbf{s}_0$ is on the left torus. If $\mathbf{s}_1$ is in the red (shaded) part, then we keep it on the left torus. If $\mathbf{s}_1$ is in the green (white) part, then we move it to the corresponding point on the right torus. In general, $\mathbf{s}_n$ is on a particular torus. If $\mathbf{s}_{n+1}$ is in the red (shaded) part,

2010 *Mathematics Subject Classification.* 11K38, 37E35.
Key words and phrases. geodesics, uniformity.

then we keep it on the same torus. If $\mathbf{s}_{n+1}$ is in the green (white) part, then we move it to the corresponding point on the other torus. Thus the sequence $\mathbf{s}_0, \mathbf{s}_1, \mathbf{s}_2, \mathbf{s}_3, \ldots$ moves from one torus to the other whenever it hits the green part. The problem is then to describe the distribution of this sequence in the union of the two tori, clearly a *parity* problem motivated by the Kronecker–Weyl equidistribution theorem.

We can visualize this discrete 2-torus system on the plane as a simple continuous system in 3-space. Figure 1.3 illustrates this observation in the case of the simpler 2-coloring in the picture on the left in Figure 1.1.

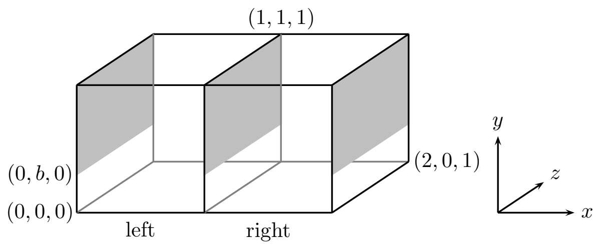

Figure 1.3: 2-cube-b 3-manifold with repeated barriers

Here there are three yz-parallel square faces of the 2-cube solid, each of which is in part a barrier, colored red (shaded), and in part a gate, colored green (white). The latter is non-empty, and permits travel between the two cubes. For these two cubes to form a 3-manifold, we have to guarantee that it is boundary-free. We use boundary identification which is a modification of the boundary identification for the torus $[0,1)^3$. The two xy-parallel square faces with $z = 0$ are identified with the two xy-parallel square faces with $z = 1$ by trivial perpendicular translation. The two xz-parallel square faces with $y = 0$ are identified with the two xz-parallel square faces with $y = 1$ by trivial perpendicular translation. The right side of the red (shaded) rectangle on the yz-parallel square face with $x = 0$ is identified with the left side of the red (shaded) rectangle on the yz-parallel square face with $x = 1$, while the right side of the red (shaded) rectangle on the yz-parallel square face with $x = 1$ is identified with the left side of the red (shaded) rectangle on the yz-parallel square face with $x = 2$. Finally, the green (white) rectangle on the yz-parallel square face with $x = 0$ is identified with the green (white) rectangle on the yz-parallel square face with $x = 2$. For convenience, we refer to this as the 2-cube-b 3-manifold.

We thus have a flat 3-manifold, with euclidean metric almost everywhere, and with boundary identification via perpendicular translation. Thus geodesic flow in this 3-manifold is 1-direction linear flow. It moves rather like 1-direction geodesic flow on the torus $[0,1)^3$, and the novelty comes from the effect of the barriers.

There is clearly an equivalence between the discrete 2-dimensional 2-torus system and this new continuous 3-dimensional 2-cube system. An infinite $\mathbf{v}_0$-shift sequence is equidistributed on the 2-torus with the 2-coloring given in the picture on the left of Figure 1.1 if and only if the corresponding half-infinite 1-direction geodesic with direction vector $\mathbf{v}$ is equidistributed in the 2-cube-b 3-manifold.

Assume now that $\mathbf{v} = (1, \alpha_1, \alpha_2) \in \mathbb{R}^3$ is a Kronecker vector. Is it true that every half-infinite 1-direction geodesic with direction vector $\mathbf{v}$ is equidistributed in any 2-cube-b 3-manifold with $0 < b < 1$?

It turns out that for every Kronecker vector $\mathbf{v}$, there are infinitely many values of the parameter b for which equidistribution fails. To explain this, we need to look at the corresponding problem in lower dimension. The projection of the 2-cube-b 3-manifold to the xy-plane gives rise to the 2-square-b surface which arises from the work of Veech [6]. Some of the anti-equidistribution results on such surfaces obtained recently by the authors in [1] can be converted to anti-equidistribution results on 2-cube-b 3-manifolds.

For instance, let $\alpha_1 \in (0, 1/2)$ be irrational, and let $b = 2\alpha_1$. Then for every α_2 for which $\mathbf{v} = (1, \alpha_1, \alpha_2) \in \mathbb{R}^3$ is a Kronecker vector, every half-infinite 1-direction geodesic with direction vector $\mathbf{v}$ violates equidistribution in the 2-cube-b 3-manifold. For more details, see [1, Theorem 2.1].

The papers [1] and [6] contain some equidistribution results on the 2-square-b surface. These, unfortunately, do not immediately lead to corresponding results on the 2-cube-b 3-manifold. Nevertheless, using a different approach, we can establish equidistribution for most half-infinite geodesics in the 2-cube-b 3-manifold. Furthermore, we can generalize the result to any 2-coloring of the unit torus $[0, 1)^2$ where each of the red and green parts is the union of finitely many polygons, and where the green part has positive area. The richness of the possibilities to fix such 2-colorings is particularly interesting.

Indeed, we can consider an n-torus system, with n copies of the unit torus $[0, 1)^2$, where $n \geqslant 2$ is an integer. This then leads to a flat 3-manifold, with euclidean metric almost everywhere, and with boundary identification via perpendicular translation. For instance, if we take $n = 4$ and use the 2-coloring of the torus $[0, 1)^2$ as shown in the picture on the right in Figure 1.1, then we have the 4-cube 3-manifold with repeated barriers as shown in Figure 1.4.

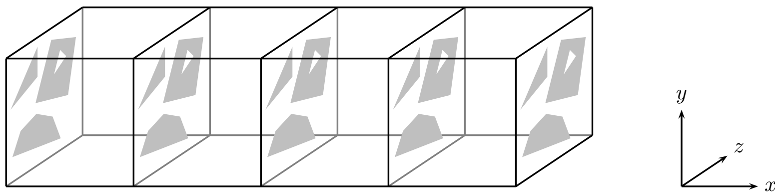

Figure 1.4: a 4-cube 3-manifold with repeated barriers

Theorem 1. *Let $n \geqslant 2$ be an integer, and let $\mathcal{M}$ be any n-cube 3-manifold with barriers, where the yz-parallel square faces have identical 2-coloring such that each of the red and green parts is the union of finitely many polygons, and where the green part has positive area. Then for almost every starting point and almost every direction $\mathbf{v} = (1, \alpha_1, \alpha_2) \in \mathbb{R}^3$, the corresponding half-infinite 1-direction geodesic is equidistributed in $\mathcal{M}$.*

As a trivial corollary, we deduce that the half-infinite 1-direction geodesic spends asymptotically the same amount of time in each one of the n cubes of the n-cube 3-manifold.

We remark that any polygon in the given 2-coloring can be replaced by a circle, an ellipse, or any other *piecewise smooth closed curve*. It requires an extra analytic discussion in the proof that we postpone to Section 6.

An immediate question that arises is whether we can extend Theorem 1 to include every Kronecker direction $\mathbf{v} \in \mathbb{R}^3$.

By a grid type 2-coloring of the torus $[0, 1)^2$, we mean dividing the torus $[0, 1)^2$ into m^2 subsquares in the natural way, where m is a positive integer, and coloring at least one of the subsquares green and the remainder red, as illustrated in Figure 1.5.

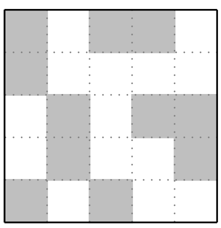

Figure 1.5: grid type 2-coloring of the torus $[0, 1)^2$

Using a similar method, the authors can establish the following stronger conclusion in this setting; see [2, Theorem 4].

Theorem 2. *Let $n \geqslant 2$ be an integer, and let $\mathcal{M}$ be any n-cube 3-manifold with barriers, where the yz-parallel square faces have identical grid type 2-coloring, and where the green part has positive area. Then every half-infinite 1-direction geodesic with a Kronecker direction $\mathbf{v} = (1, \alpha_1, \alpha_2) \in \mathbb{R}^3$ is equidistributed in $\mathcal{M}$.*

Figure 1.6 shows the 2-cube box with a 2-coloring on the middle yz-parallel square face such that each of the red (shaded) and green (white) parts is the union of finitely many polygons, and where the green part has positive area. Consider billiard in this 2-cube box, where in addition to the square faces on the surface of the box, there are additional barriers on the middle yz-parallel square face given by the parts colored red. As usual, we consider the ideal case of a point billiard that bounces back at any barrier, following the well-known rules of optical reflection.

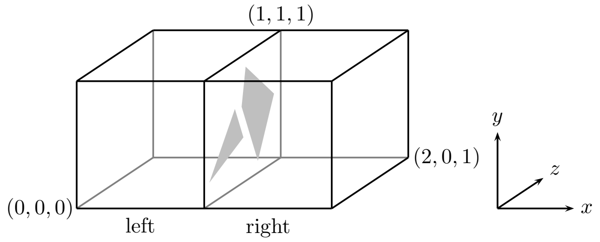

Figure 1.6: 2-cube box with barriers in the middle

We are interested in the long term bahavior of the billiard orbit. We shall show that Theorem 1 contributes to our understanding of such questions.

To establish equidistribution for such billiard orbits, we extend the idea of König and Szücs [4] and apply 3-dimensional *unfolding*. This converts the billiard orbit in this 2-cube box with barriers into a 1-direction geodesic in a boundary-free flat 3-manifold. The latter system is an 8-copy construction involving 16 cubes, and results from three consecutive reflections across a plane.

The original 2-cube box with barriers in the middle is highlighted in bold in Figure 1.7. We reflect it across the plane $x = 2$, then reflect the 2-copy union across the plane $y = 1$, and finally reflect the 4-copy union across the plane $z = 1$ to obtain an 8-copy union. Thus the original $2 \times 1 \times 1$ box becomes a $4 \times 2 \times 2$ box with two repeated sets of barriers on the yz-parallel squares $[0, 2)^2$ on the faces $x = 1$ and $x = 3$.

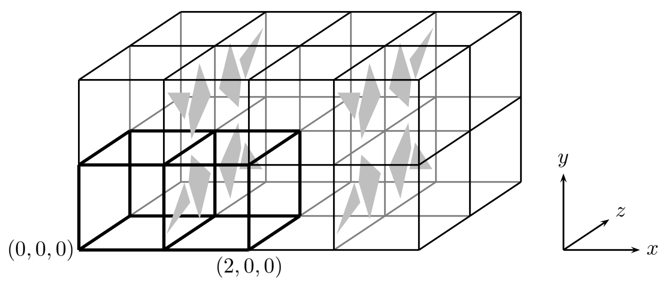

Figure 1.7: unfolding the billiard orbit in a 2-cube box
with barriers in the middle

This box has boundary, and we turn it into a boundary-free flat 3-manifold with boundary identification. First of all, the square faces on the boundary of this box are identified by perpendicular translation. Next, let L and R denote respectively the barriers on the square faces on the plane $x = 1$ and $x = 3$ respectively. The left side of L is identified with the right side of R, while the right side of L is identified with the left side of R. For convenience, we refer to this special 3-manifold as the *2-cube-billiard 3-manifold*.

Clearly 1-direction geodesic flow in the 2-cube-billiard 3-manifold is an 8-fold cover of billiard flow in the 2-cube box with barriers.

We now remove the part $[0,1) \times [0,2) \times [0,2)$ on the left and join it instead to the right to become $[4,5) \times [0,2) \times [0,2)$, as shown in Figure 1.8.

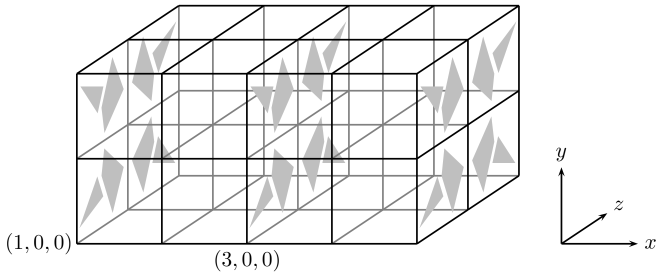

Figure 1.8: cutting and pasting

Contracting the resulting 3-manifold by a factor 1/2 in each of the three directions, we then obtain a 2-cube 3-manifold with repeated barriers that we have studied in Theorem 1. The following result is then a corollary of Theorem 1 in the special case $n = 2$.

Theorem 3. *Consider billiard in a special 2-cube box with barriers in the middle square face joining the cubes given by a 2-coloring such that each of the red and green parts is the union of finitely many polygons, and where the green part has positive area. Then for almost every starting point and almost every initial direction* $\mathbf{v} = (1, \alpha_1, \alpha_2) \in \mathbb{R}^3$ *with* $\mathbf{v}_0 = (\alpha_1, \alpha_2) \in [-1,1]^2$, *the corresponding half-infinite billiard orbit is equidistributed in this special 2-cube box.*

Unfortunately, Theorem 1 does not seem to help in the case of more complicated billiards with barriers.

Of course, there is no reason why the 2-coloring on the distinct yz-parallel square faces of the n-cube 3-manifold should be the same, apart from possibly making the problem a little simpler.

We know that if a geodesic hits an xy-parallel square face of $\mathcal{M}$, then it jumps to the corresponding point on the identified xy-parallel square face and continues in the same direction, and if a geodesic hits an xz-parallel square face, then it jumps to the corresponding point on the identified xz-parallel square face and continues in the same direction.

Suppose now that a geodesic with direction $\mathbf{v} = (1, \alpha_1, \alpha_2)$ hits a yz-parallel square face at a point P. Then the continuation of the geodesic depends on the coloring of the intersection point P. If P is green, then the geodesic continues on its way in the same direction. If P is red, then we consider a directed line starting from P in the direction $(-1, 0, 0)$. This line will hit a red point P' for the first time. Then the geodesic continues from P' in the same direction, as shown in Figure 1.9.

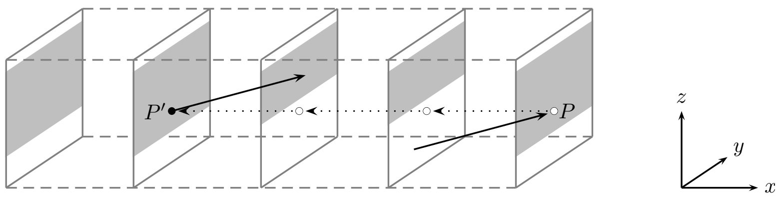

Figure 1.9: when a geodesic hits a red barrier

There is the pathological case that $P' = P$, so that the geodesic continues on its way as if P were green. To avoid such situations, we deem the point P to be colored green. This is some kind of automatic recoloring, but the following rule is a little simpler.

Restriction on Red Coloring. On any line perpendicular to any given yz-parallel square face of $\mathcal{M}$, there are either no points colored red or at least 2 distinct points colored red. Here we use the standard convention that the left-most yz-parallel square face and the right-most yz-parallel square face are the same.

Instead of requiring perfect repetition of the 2-coloring on all the yz-parallel square faces as in Theorem 1, we now impose the substantially weaker condition of *local repetition*. More precisely, we require a small *local repetition color-split neighborhood*, in the form of a line segment with the same local 2-coloring of red and green in the two opposite side-neighborhoods. For illustration, see Figure 1.10, where the three highlighted rectangles are in the same position within the square torus.

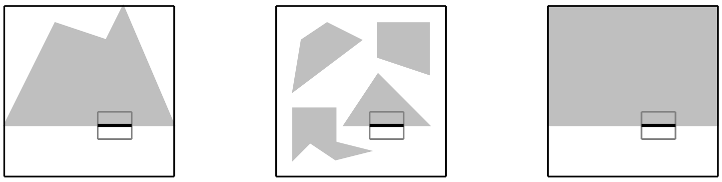

Figure 1.10: different 2-colorings with local repetition

We emphasize that the local repetition color-split neighborhood must be present on all the yz-parallel square faces of the n-cube 3-manifold. Naturally, we still need 2-colorings on each yz-parallel square face such that each of the red and green parts is the union of finitely many polygons, and where the green part has positive area. Furthermore, we also require that the Restriction on Red Coloring holds. Since the 2-colorings on the different yz-parallel squares faces can now be different, this represents substantially more freedom for the 2-colorings. In Figure 1.11, we have a 4-cube 3-manifold with local repetition color-split provided by the triangular red (shaded) regions at the corners of the yz-parallel squares faces. The positions of the local repetition color-split neighborhood on the different yz-parallel squares faces are indicated by the short thick lines.

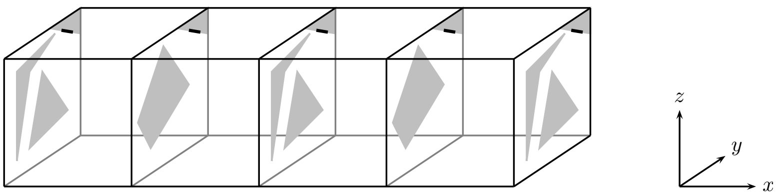

Figure 1.11: a 4-cube 3-manifold with local repetition

The result that we can prove in this more general setting is not really weaker.

Theorem 4. *Let $n \geqslant 2$ be an integer, and let $\mathcal{M}$ be any n-cube 3-manifold with barriers, where each yz-parallel square face has a 2-coloring such that each of the red and green parts is the union of finitely many polygons, and where the green part has positive area. Suppose further that the Restriction on Red Coloring holds, and that there is a local repetition color-split neighborhood on the yz-parallel square faces. Then for almost every starting point and almost every direction $\mathbf{v} = (1, \alpha_1, \alpha_2) \in \mathbb{R}^3$, the corresponding half-infinite 1-direction geodesic is equidistributed in $\mathcal{M}$.*

As in Theorem 1, any polygon of the 2-colorings can be replaced by a circle, an ellipse, or any other piecewise smooth closed curve.

2. Preparation for the proof of Theorem 4

We work with an equivalent discrete form of the problem.

For every integer $i = 0, 1, \ldots, n-1$, let

$$U_i = \{i\} \times [0,1)^2$$

denote the i-th yz-parallel square face of $\mathcal{M}$, and let

$$X_0 = \bigcup_{i=0}^{n-1} U_i.$$

Since each yz-parallel square face has a 2-coloring, for every integer $i = 0, 1, \ldots, n-1$, there exist two sets $\mathscr{G}_i$ and $\mathscr{R}_i$ such that

$$U_i = \mathscr{G}_i \cup \mathscr{R}_i \quad \text{and} \quad \mathscr{G}_i \cap \mathscr{R}_i = \emptyset. \tag{2.1}$$

Since the collection of non-Kronecker vectors $\mathbf{v} = (1, \alpha_1, \alpha_2) \in \mathbb{R}^3$ has measure zero, we may therefore start our discussion by assuming that $\mathbf{v}$ is a Kronecker vector. Let $\mathbf{v} = (1, \alpha_1, \alpha_2) \in \mathbb{R}^3$ be a Kronecker vector, and let $\mathbf{v}_0 = (\alpha_1, \alpha_2)$. We define an invertible transformation $\mathbf{T} = \mathbf{T}_{\mathbf{v}} : X_0 \to X_0$ as follows.

For any point $P = (i, \mathbf{y}) \in U_i \subset X_0$ where $\mathbf{y} = (y, z) \in [0,1)^2$, let

$$\mathbf{T}(P) = \begin{cases} (i+1, \{\mathbf{y} + \mathbf{v}_0\}), & \text{if } P + \mathbf{v} = (i+1, \{\mathbf{y} + \mathbf{v}_0\}) \in \mathscr{G}_{i+1}, \\ (i^*, \{\mathbf{y} + \mathbf{v}_0\}), & \text{if } P + \mathbf{v} = (i+1, \{\mathbf{y} + \mathbf{v}_0\}) \in \mathscr{R}_{i+1}, \end{cases} \tag{2.2}$$

where the addition in $P + \mathbf{v}$ is modulo n for the first coordinate and modulo 1 for the remaining coordinates. Thus

$$\{\mathbf{y} + \mathbf{v}_0\} = (\{y + \alpha_1\}, \{z + \alpha_2\}), \tag{2.3}$$

where $0 \leqslant \{\beta\} < 1$ denotes the fractional part of a real number β. Furthermore, the value of i^* is determined by

$$i^* = \max\{j < i+1 : (j, \{\mathbf{y} + \mathbf{v}_0\}) \in \mathscr{R}_j\}, \tag{2.4}$$

with the convention that

$$i+1 > i > i-1 > \ldots > 1 > 0 > n-1 > n-2 > \ldots > i+2. \tag{2.5}$$

Note that (2.4) and (2.5) are motivated by Figure 1.9, and Figure 2.1 illustrates the special case when $n = 4$. If $P + \mathbf{v}$ is a red point, we then move from $P + \mathbf{v}$ in the direction of the vector $(-1, 0, 0)$ until we hit the first red point, and this red point lies on the yz-parallel square face U_{i^*}.

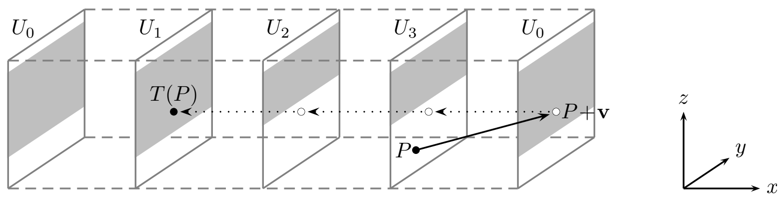

Figure 2.1: when $P + \mathbf{v}$ is red

Projecting the transformation $\mathbf{T}$ by ignoring the first coordinate, we obtain the invertible transformation

$$\mathbf{T}_0 = \mathbf{T}_{\mathbf{v}_0} : [0,1)^2 \to [0,1)^2 : \mathbf{y} \mapsto \{\mathbf{y} + \mathbf{v}_0\}.$$

It is easy to see that $\mathbf{T}$ preserves the 2-dimensional Lebesgue measure λ_2. Our goal is to establish that for almost every vector $\mathbf{v}_0 \in [0,1]^2$, the transformation $\mathbf{T} = \mathbf{T}_{\mathbf{v}} : X_0 \to X_0$ is ergodic, where $\mathbf{v} = (1, \mathbf{v}_0)$. The basic idea is quite surprising, as we prove ergodicity for this non-integrable system by taking advantage of the split singularities.

Remark. We are going to use Birkhoff's well known pointwise ergodic theorem on measure preserving transformations. Since we simply apply ergodic theory, we do not expect the reader to have any serious expertise in the subject. Thus knowledge of Lebesgue integral and basic measure theory suffices. The theorem concerns a measure-preserving system $(X, \mathcal{A}, \mu, \mathbf{T})$. Here $(X, \mathcal{A}, \mu)$ is a measure space, where X is the underlying space, $\mathcal{A}$ is a σ-algebra of subsets of X and μ is a non-negative σ-additive measure on X with $\mu(X) < \infty$, while $\mathbf{T} : X \to X$ is a measurable map which is measure-preserving, so that $\mathbf{T}^{-1}A \in \mathcal{A}$ and $\mu(\mathbf{T}^{-1}A) = \mu(A)$ for every $A \in \mathcal{A}$.

Let $L^1(X, \mathcal{A}, \mu)$ denote the space of measurable and integrable functions in the measure space $(X, \mathcal{A}, \mu)$. Then the general form of Birkhoff's pointwise ergodic theorem says that for every function $f \in L^1(X, \mathcal{A}, \mu)$, the limit

$$\lim_{m\to\infty} \frac{1}{m} \sum_{j=0}^{m-1} f(\mathbf{T}^j \mathbf{x}) = f^*(\mathbf{x}) \tag{2.6}$$

exists for μ-almost every $\mathbf{x} \in X$, where $f^* \in L^1(X, \mathcal{A}, \mu)$ is a $\mathbf{T}$-invariant measurable function satisfying the condition

$$\int_X f \, \mathrm{d}\mu = \int_X f^* \, \mathrm{d}\mu.$$

A particularly important special case is if $\mathbf{T}$ is *ergodic*, when every measurable $\mathbf{T}$-invariant set $A \in \mathcal{A}$ is *trivial* in the precise sense that $\mu(A) = 0$ or $\mu(A) = \mu(X)$. This is equivalent to the assertion that every measurable $\mathbf{T}$-invariant function is constant μ-almost everywhere.

If $\mathbf{T}$ is ergodic, then (2.6) simplifies to

$$\lim_{m\to\infty} \frac{1}{m} \sum_{j=0}^{m-1} f(\mathbf{T}^j \mathbf{x}) = \int_X f \, \mathrm{d}\mu, \tag{2.7}$$

and the right-hand side of (2.6) is the same constant for μ-almost every $\mathbf{x} \in X$.

The remarkable intuitive interpretation of (2.7) is that the *time average* on the left hand side is equal to the *space average* on the right hand side.

3. Proof of Theorem 4

We focus on the particular measure-preserving system $(X_0, \mathcal{A}, \lambda_2, \mathbf{T})$, where $\mathcal{A}$ is the family of Borel sets in X_0, λ_2 is 2-dimensional Lebesgue measure and $\mathbf{T} = \mathbf{T}_{\mathbf{v}}$. We shall establish ergodicity by contradiction.

Step 1. Suppose on the contrary that $\mathbf{T}$ is not ergodic. Then there exists a non-trivial measurable $\mathbf{T}$-invariant subset $S_0 \subset X_0$ such that $0 < \lambda_2(S_0) < n$. We try to derive a contradiction.

Removing possibly a set of λ_2-measure zero, we may assume that for every point $\mathbf{x} \in X_0$, the point $\mathbf{T}^j(\mathbf{x})$ is well defined for every integer $j = 1, 2, 3, \ldots$.

Lemma 3.1. *Consider the measure-preserving system $(X_0, \mathcal{A}, \lambda_2, \mathbf{T})$, where $\mathbf{T} = \mathbf{T}_{\mathbf{v}}$ and $\mathbf{v} = (1, \mathbf{v}_0)$ is a Kronecker vector. For any $\mathbf{T}$-invariant subset $S_0 \subset X_0$, let the multiplicity function $\widetilde{\chi}_{S_0}$ of S_0 be defined for every point $P \in [0,1)^2$ by*

$$\widetilde{\chi}_{S_0}(P) = |\{i = 0, 1, \ldots, n-1 : (i, P) \in S_0\}|.$$

Suppose further that S_0 is a proper subset of X_0, so that $S_0 \neq \emptyset$ and $S_0 \neq X_0$. Then there exists an integer $k_0 = 1, \ldots, n-1$ such that $\widetilde{\chi}_{S_0}(P) = k_0$ for almost every point $P \in [0,1)^2$, so that $\lambda_2(S_0) = k_0$.

Proof. Since $\mathbf{v}$ is a Kronecker vector, it follows that the $\mathbf{v}_0$-shift on the unit torus $[0,1)^2$ is ergodic. Meanwhile, it is easy to check that the multiplicity function $\widetilde{\chi}_{S_0}$ is $\mathbf{T}_0$-invariant. Thus Birkhoff's ergodic theorem implies that $\widetilde{\chi}_{S_0}$ is constant almost everywhere. Note that $\widetilde{\chi}_{S_0}$ is integer valued and cannot be equal to 0 or n. This completes the proof. □

Step 2. Given a point $\mathbf{z} \in X_0$ and a radius $0 < r < 1/2$, let $D(\mathbf{z};r)$ denote the circular disk of radius r and center $\mathbf{z}$. Clearly $D(\mathbf{z};r)$ has area πr^2. Note that $D(\mathbf{z};r) \subset X_0$, due to the fact that X_0 is a compact flat surface.

Since the non-trivial $\mathbf{T}$-invariant subset $S_0 \subset X_0$ is measurable, it follows from Lebesgue's density theorem that for almost every $\mathbf{z} \in S_0$,

$$\lim_{r\to 0} \frac{\lambda_2(S_0 \cap D(\mathbf{z};r))}{\pi r^2} = 1,$$

whereas for almost every $\mathbf{z} \in S_0^c = X_0 \setminus S_0$,

$$\lim_{r\to 0} \frac{\lambda_2(S_0 \cap D(\mathbf{z};r))}{\pi r^2} = 0.$$

Let M be a large integer, and divide each of $U_0, U_1, \ldots, U_{n-1}$ into M^2 congruent squares of area $(1/M)^2$ in the standard way. We refer to these small squares as special $(1/M)$-squares. Thus there are precisely nM^2 special $(1/M)$-squares in X_0.

In view of Lebesgue's density theorem, we formulate here and prove in Section 4 the following lemma for the hypothetical non-trivial measurable $\mathbf{T}$-invariant subset $S_0 \subset X_0$.

Lemma 3.2. *Let the real number $\varepsilon \in (0,1)$ be arbitrarily small and fixed, and let the real number $\varepsilon_1 > 0$ be fixed. There exists a finite threshold $m_0 = m_0(S_0;\varepsilon;\varepsilon_1)$ such that for every integer $M \geqslant m_0$, there exist at least $(1-\varepsilon_1)nM^2$ special $(1/M)$-squares Q in X_0 such that either*

$$\frac{\lambda_2(S_0 \cap Q)}{(1/M)^2} > 1 - \varepsilon \quad \textit{or} \quad \frac{\lambda_2(S_0 \cap Q)}{(1/M)^2} < \varepsilon.$$

Let N be a large even integer. Let $\mathcal{F}(N/2)$ denote the standard decomposition of the unit torus $[0,1)^2$ into $(N/2)^2$ axis-parallel congruent small squares of common side length $2/N$ such that the origin $(0,0)$ is the vertex of a small square. For $\boldsymbol{\delta} = (\delta_1,\delta_2) \in \{0,1\}^2$, let $\mathcal{F}_{\boldsymbol{\delta}}(N/2)$ denote the translation of $\mathcal{F}(N/2)$ modulo one such that the vertex $(0,0)$ moves to $(\delta_1/N, \delta_2/N)$. We refer to the small squares in the four partitions $\mathcal{F}_{\boldsymbol{\delta}}(N/2)$, $\boldsymbol{\delta} \in \{0,1\}^2$, as basic $(2/N)$-squares. It is not difficult to see that any axis-parallel square B of side length $1/N$ in the unit torus is contained in a basic $(2/N)$-square.

For every $i = 0,1,\ldots,n-1$, we replicate the families $\mathcal{F}_{\boldsymbol{\delta}}(N/2)$, $\boldsymbol{\delta} \in \{0,1\}^2$, on the yz-parallel square face $U_i = \{i\} \times [0,1)^2$. In other words, we write

$$\mathcal{F}_{\boldsymbol{\delta}}(N/2;i) = \{i\} \times \mathcal{F}_{\boldsymbol{\delta}}(N/2), \quad \boldsymbol{\delta} \in \{0,1\}^2, \quad i = 0,1,\ldots,n-1.$$

For every $\boldsymbol{\delta} \in \{0,1\}^2$, Lemma 3.2 with $M = N/2$ then gives the following.

Let the real number $\varepsilon \in (0,1)$ be arbitrarily small and fixed, and let the real number $\varepsilon_1 > 0$ be fixed. There exists a finite threshold $m_0 = m_0(S_0;\varepsilon;\varepsilon_1)$ such that for every $\boldsymbol{\delta} \in \{0,1\}^2$ and every even integer $N \geqslant m_0$, there exist at least $(1-\varepsilon_1)n(N/2)^2$ basic $(2/N)$-squares

$$Q \in \bigcup_{i=0}^{n-1} \mathcal{F}_{\boldsymbol{\delta}}(N/2;i) \tag{3.1}$$

such that either

$$\frac{\lambda_2(S_0 \cap Q)}{(2/N)^2} > 1 - \varepsilon \quad \text{or} \quad \frac{\lambda_2(S_0 \cap Q)}{(2/N)^2} < \varepsilon.$$

We say that these sets Q satisfy the *ε-nearly zero-one law.* The conclusion of this step is that we have a lower bound for the number of such basic $(2/N)$-squares of the form (3.1).

Step 3. To apply a version of the splitting method, first introduced in [2], we need to make use of the local repetition color-split neighborhood $\mathcal{S}$, with green part $\mathcal{S}_{\mathscr{G}}$ and red part $\mathcal{S}_{\mathscr{R}}$. As this color-split neighborhood is present on all the yz-parallel square faces of the n-cube 3-manifold $\mathcal{M}$, we assume, for simplicity, that $\mathcal{S}$ lies on the unit torus $[0,1)^2$.

We can clearly find within $\mathcal{S}$ a buffer zone $\mathcal{B}_N$ as shown in Figure 3.1. The length of $\mathcal{B}_N$ is a constant c_1 depending only in $\mathcal{S}$, while the width is $1/2N$, where the integer N is sufficiently large, with the color-split boundary splitting $\mathcal{B}_N$ into green and red strips of width $1/4N$.

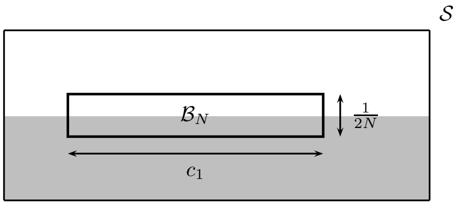

Figure 3.1: buffer zone $\mathcal{B}_N$ within the local repetition color-split neighborhood $\mathcal{S}$

The following result follows from simple geometric considerations.

Lemma 3.3. *Let $A_0 \subset [0,1)^2$ be an arbitrary axis-parallel square with side length $1/N$ and center $c(A_0)$. Then for every $\mathbf{t} \in [0,1)^2$ such that $c(A_0) + \mathbf{t} \in \mathcal{B}_N$, the set $A_0 + \mathbf{t}$ has substantial color-split in the sense that*

$$\lambda_2((A_0 + \mathbf{t}) \cap \mathcal{S}_{\mathscr{G}}) \geqslant \frac{1}{14N^2} \quad \text{and} \quad \lambda_2((A_0 + \mathbf{t}) \cap \mathcal{S}_{\mathscr{R}}) \geqslant \frac{1}{14N^2}.$$

The extreme case of the above takes place when $\mathcal{B}_N$ is tilted at 45 degrees, with the center $c(A_0)$ of A_0 at the corner, as shown in Figure 3.2.

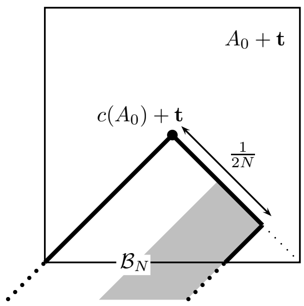

Figure 3.2: extreme case of substantial color-split

We need to show that such substantial color-splits occur quite frequently. We shall establish in Section 5 the following result.

Lemma 3.4. *Suppose that the integer N is even and sufficiently large. For every pair $\mathbf{s}_0 \in [0,1)^2$ and $\mathbf{v}_0 \in [0,1]^2$, let $F(\mathbf{s}_0; \mathbf{v}_0; N^2)$ denote the number of integers $j = 0, 1, \ldots, N^2 - 1$ such that $\mathbf{s}_0 + j\mathbf{v}_0 \in \mathcal{B}_N$. Then for every $\varepsilon_2 > 0$, there exists a constant $c_2 > 0$ such that*

$$\lambda_2(\{\mathbf{v}_0 \in [0,1)^2 : F(\mathbf{s}_0; \mathbf{v}_0; N^2) \geqslant c_2 N\}) \geqslant 1 - \varepsilon_2. \tag{3.2}$$

Step 4. Consider a $(1/N)$-square A on some yz-parallel square face U_i of $\mathcal{M}$. We see from (2.1)–(2.5) that the image $\mathbf{T}(A)$ is then given by $A+\mathbf{v}$ followed by appropriate bounce-backs on those parts that hit red. If $A+\mathbf{v}$ hits both green and red, then clearly the splitting of the image $\mathbf{T}(A)$ is caused by a *color-split*. For instance, in Figure 3.3, the white square on the yz-parallel square face U_1 denotes A, the square $A+\mathbf{v}$ experiences a color-split on the yz-parallel square face U_2, and the image $\mathbf{T}(A)$, indicated in black, is split between the yz-parallel square faces U_1 and U_2. Indeed, if the 2-coloring on the different yz-parallel square faces of $\mathcal{M}$ are identical, then the only splitting of image is caused by a color-split.

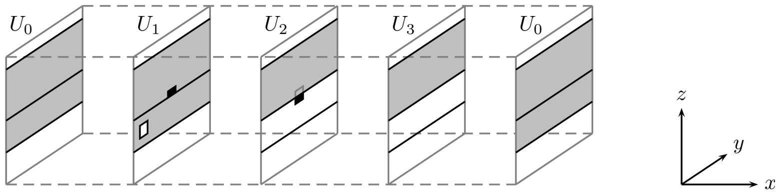

Figure 3.3: a color-split

On the other hand, if the different yz-parallel square faces of $\mathcal{M}$ can have different 2-colorings, then there are other instances that cause splitting of the image $\mathbf{T}(A)$. For instance, in Figure 3.4, for the white square A on the yz-parallel square face U_3, the square $A+\mathbf{v}$ is on the yz-parallel square face U_0, but then different parts have different bounce-backs, and the image $\mathbf{T}(A)$ is split between the yz-parallel square faces U_1 and U_3. Although a color-split always leads to an image-split, this example shows that the converse is not true.

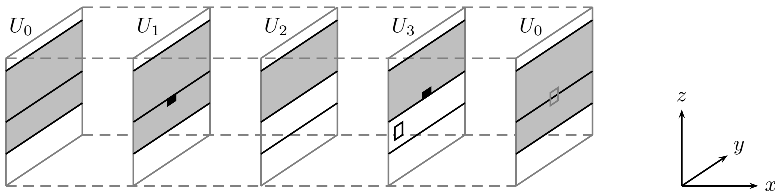

Figure 3.4: an image-split that is not a color-split

It turns out that it is advantageous to consider another kind of splitting which may or may not result in an image-split. For every $i=0,1,\ldots,n-1$, let $\{i\}\times\Gamma_i$ denote the color-split boundaries of the 2-coloring on U_i, so that $\Gamma_i\subset[0,1)^2$ is a collection of straight edges. Consider the union

$$\Gamma=\bigcup_{i=0}^{n-1}\Gamma_i,$$

and, for every $i=0,1,\ldots,n-1$, replace $\{i\}\times\Gamma_i$ on U_i by $\{i\}\times\Gamma$, so that the different yz-parallel square faces of $\mathcal{M}$ now have identical Γ-*splits*. Although an image-split is always a Γ-split, the example in Figure 3.5 shows that the converse is not true. Here, and also in Figures 3.3 and 3.4, the Γ-splits are indicated by the black line segments.

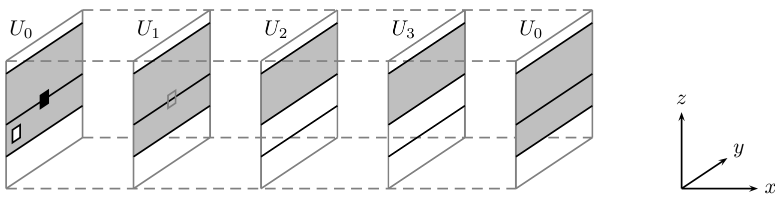

Figure 3.5: a Γ-split that is not an image-split

The important point here is that if there is no Γ-split, then there is no splitting of any kind and in particular no color-split.

For the splitting method to work, we need a substantial color-split in the local repetition color-split neighborhood and a *long* color-split-free chain either side of it. In view of the above observation, it therefore suffices to find a substantial color-split in the local repetition color-split neighborhood and a *long* Γ-split-free chain either side of it. As the Γ-split is identical on all the yz-parallel square faces of $\mathcal{M}$, this allows us to consider $\mathbf{v}_0$-flow on the unit torus $[0,1)^2$ instead of geodesic flow in the direction $\mathbf{v}$ in $\mathcal{M}$.

To establish this, we shall show that the number of *short* Γ-split-free chains is *small*. More precisely, we have the following lemma which we prove in Section 4. Before we can state the result, we first need some definitions.

Let $A_0 \subset [0,1)^2$ be an arbitrary axis-parallel square of side length $1/N$. Consider the N^2 $\mathbf{v}_0$-shift images

$$A_0 + j\mathbf{v}_0, \quad j = 0, 1, \ldots, N^2 - 1, \tag{3.3}$$

in $[0,1)^2$. We say that a set $\{A_0 + j\mathbf{v}_0 : j \in J\}$, where J is a subset of consecutive integers in $\{0, 1, \ldots, N^2 - 1\}$, is Γ-split-free if $(A_0 + j\mathbf{v}_0) \cap \Gamma = \emptyset$ for every $j \in J$. Furthermore, we say that the set $\{A_0 + j\mathbf{v}_0 : j \in J\}$ is a Γ-split-free chain if it is Γ-split-free and not contained in a bigger Γ-split-free set. It is convenient to define the length of the Γ-split-free chain to be $|J| + 1$.

Thus the sequence (3.3) decomposes into a subsequence of Γ-split members, with any two consecutive members of this subsequence possibly separated by a Γ-split-free chain in between.

For every axis-parallel square $A_0 \subset [0,1)^2$ of side length $1/N$, we can clearly write $A_0 = A_0(\mathbf{s}_0)$, where $\mathbf{s}_0 = c(A_0)$ is the center of A_0. Also, for every vector $\mathbf{v} = (1, \mathbf{v}_0)$, where $\mathbf{v}_0 \in \mathbb{R}^2$, we can identify

$$\mathbf{v}, \quad \mathbf{v}_0 \quad \text{and} \quad \mathbf{v}^* = \frac{\mathbf{v}}{|\mathbf{v}|}$$

with each other. Clearly the collection of all vectors $\mathbf{v}^*$, where $\mathbf{v}_0 \in \mathbb{R}^2$, forms the subset $\mathbf{S}^2_+ = \{(x,y,z) \in \mathbf{S}^2 : x > 0\}$ of the unit ball $\mathbf{S}^2$.

Lemma 3.5. *Let the real number $\varepsilon_2 \in (0,1)$ be arbitrarily small and fixed, and let $A_0 \subset [0,1)^2$ be an arbitrary axis-parallel square of side length $1/N$ that runs uniformly over $[0,1)^2$, in the sense that the center $\mathbf{s}_0$ is uniformly distributed on $[0,1)^2$. Then there exists a positive absolute constant c_3 such that for at least $(1-\varepsilon_2)$-proportion of the pairs*

$$(\mathbf{s}_0, \mathbf{v}^*) \in [0,1)^2 \times \mathbf{S}^2_+,$$

the sequence (3.3) in $[0,1)^2$ contains at most $\varepsilon_2 N$ Γ-split-free chains with length at most $c_3 \varepsilon_2^2 N$. The $(1-\varepsilon_2)$-proportion is in terms of the product of the 2-dimensional Lebesgue measure on $[0,1)^2$ and the normalized surface area measure on $\mathbf{S}^2_+$.

Step 5. The motivation for this step is not immediately obvious at this stage, but we include it here for the sake of convenience. This will become clear in Step 7.

We establish in Section 4 the following result on *clustering*. It demonstrates that the overwhelming majority of arithmetic progressions are not clustered.

Lemma 3.6. *Let the real number $\varepsilon_2 \in (0,1)$ be arbitrarily small and fixed. There exists a finite constant $C^* = C^*(\varepsilon_2)$ such that for any starting point $\mathbf{s}_0$ in the unit torus $[0,1)^2$ and for at least $(1-\varepsilon_2)$-proportion of the vectors $\mathbf{v}_0 \in [0,1]^2$, every axis-parallel square Q of side length $2/N$ in $[0,1)^2$ contains modulo one at most C^* elements of the arithmetic progression*

$$\mathbf{s}_0 + j\mathbf{v}_0, \quad j = 0, 1, \ldots, N^2 - 1,$$

of N^2 terms. In particular, we can take $C^* = 1 + 16/\varepsilon_2$.

Step 6. Here we combine our observations in Steps 3–5. Combining Lemmas 3.3, 3.4, 3.5 and 3.6, we see that there exists a set $\mathbf{GD} \subset [0,1]^2$ of *good* directions $\mathbf{v}_0 \in [0,1]^2$ with $\lambda_2(\mathbf{GD}) \geqslant 1 - 3\varepsilon_2$ such that for every $\mathbf{v}_0 \in \mathbf{GD}$, the conclusion of Lemma 3.6 holds for any $\mathbf{s}_0 \in [0,1)^2$, and there exists an axis-parallel square A_0^* in $[0,1)^2$ of side length $1/N$ such that the following two conditions hold:

(i) The sequence

$$A_0^* + j\mathbf{v}_0, \quad j = 0, 1, \ldots, N^2 - 1, \tag{3.4}$$

in $[0,1)^2$ contains at least c_2N members which exhibit substantial color-split in the local repetition color-split neighborhood $\mathcal{S}$, in the sense that

$$\lambda_2((A_0^* + j\mathbf{v}_0) \cap \mathcal{S}_{\mathscr{G}}) \geqslant \frac{1}{14N^2} \quad \text{and} \quad \lambda_2((A_0^* + j\mathbf{v}_0) \cap \mathcal{S}_{\mathscr{R}}) \geqslant \frac{1}{14N^2}.$$

(ii) The sequence (3.4) has at most $\varepsilon_2 N$ *short* Γ-split-free chains with length at most $c_3\varepsilon_2^2 N$.

Combining (i) and (ii), we conclude that the sequence (3.4) has at least

$$\frac{c_2N - 2\varepsilon_2 N - 2}{2} \geqslant \frac{c_2N}{3} \tag{3.5}$$

pairs of consecutive *long* Γ-split-free chains of length at least $c_3\varepsilon_2^2N$, where each pair is separated by a member of the sequence with substantial color-split in the local repetition color-split neighborhood $\mathcal{S}$. Here we require that $\varepsilon_2 > 0$ is chosen to be arbitrarily small so that the inequality (3.5) holds.

Step 7. To facilitate the use of the splitting method, we need to first demonstrate the existence of a pair of consecutive long Γ-split-free chains separated by $A_0^* + j_0\mathbf{v}_0$ with substantial color-split in the local repetition color-split neighborhood $\mathcal{S}$ and such that both $A_0^* + (j_0-1)\mathbf{v}_0$ and $A_0^* + (j_0+1)\mathbf{v}_0$ are Γ-split-free and satisfy the ε-nearly zero-one law. More precisely, we return to the n-cube 3-manifold $\mathcal{M}$. Then $A_0^* + j_0\mathbf{v}_0$ corresponds to an axis-parallel square A^{**} of side length $1/N$ on some yz-parallel square face U_i, while $A_0^* + (j_0-1)\mathbf{v}_0$ and $A_0^* + (j_0+1)\mathbf{v}_0$ correspond to the squares $A^{**} - \mathbf{v}$ and $A^{**} + \mathbf{v}$ respectively. Then we claim that either

$$\frac{\lambda_2(S_0 \cap (A^{**} - \mathbf{v}))}{\lambda_2(A^{**} - \mathbf{v})} > 1 - \varepsilon \quad \text{or} \quad \frac{\lambda_2(S_0 \cap (A^{**} - \mathbf{v}))}{\lambda_2(A^{**} - \mathbf{v})} < \varepsilon, \tag{3.6}$$

and either

$$\frac{\lambda_2(S_0 \cap (A^{**} + \mathbf{v}))}{\lambda_2(A^{**} + \mathbf{v})} > 1 - \varepsilon \quad \text{or} \quad \frac{\lambda_2(S_0 \cap (A^{**} + \mathbf{v}))}{\lambda_2(A^{**} + \mathbf{v})} < \varepsilon. \tag{3.7}$$

To establish our claim, suppose, on the contrary, that for every pair of consecutive long Γ-split-free chains separated by $A_0^* + j_0\mathbf{v}_0$ with substantial color-split in the local repetition color-split neighborhood $\mathcal{S}$, either (3.6) or (3.7) fails. Without loss of generality, suppose that (3.7) fails, so that $A^{**} + \mathbf{v}$ fails the ε-nearly zero-one law. Note that $A^{**} + \mathbf{v}$ corresponds to the first member of a Γ-split-free chain of at least $c_3\varepsilon_2^2N - 2$ terms. Since the subset $S_0 \subset X_0$ is $\mathbf{T}$-invariant, it follows that every set

$$\mathbf{T}^j(A^{**} + \mathbf{v}), \quad 0 \leqslant j \leqslant c_3\varepsilon_2^2N - 3, \tag{3.8}$$

is an axis-parallel square of side length $1/N$ on a yz-parallel square face of $\mathcal{M}$ that is Γ-split-free and fails the ε-nearly zero-one law. Taking into consideration all pairs of consecutive long Γ-split-free chains, we see that there are at least

$$\frac{c_2N}{3} \cdot (c_3\varepsilon_2^2N - 2)$$

such axis-parallel squares of the form (3.8) which fail the ε-nearly zero-one law. Each is contained in a basic $(2/N)$-square Q of the form (3.1) for some $\boldsymbol{\delta} \in \{0,1\}^2$.

On the other hand, in view of Lemma 3.6, any basic $(2/N)$-square Q contains at most $C^* = C^*(\varepsilon_2)$ such axis-parallel squares of the form (3.8) which fail the ε-nearly zero-one law. We conclude therefore that there are at least

$$\frac{c_2 N(c_3\varepsilon_2^2 N - 2)}{3C^*}$$

distinct basic $(2/N)$-squares Q of the form (3.1) for some $\boldsymbol{\delta} \in \{0,1\}^2$ which fail the ε-nearly zero-one law. Meanwhile, the conclusion from Step 2 is that the number of such basic $(2/N)$-squares Q is bounded above by

$$4\varepsilon_1 n \left(\frac{N}{2}\right)^2.$$

For any given $\varepsilon_2 > 0$, we can now choose $\varepsilon_1 > 0$ sufficiently small in terms of the other parameters to obtain a contradiction. This establishes the claim.

Step 8. We now apply the splitting method. Assume that A^{**} is an axis-parallel square of side length $1/N$ on some yz-parallel square face of $\mathcal{M}$, and that (3.6) and (3.7) both hold. For every $i = 0, 1, \ldots, n-1$, let $A_i \subset U_i$ be the axis-parallel square of side length $1/N$ that is in the same general position as A^{**}, and let

$$B_{i-1} = A_i - \mathbf{v} \quad \text{and} \quad C_{i+1} = A_i + \mathbf{v}.$$

Note that each A_i falls within the local repetition color-split neignborhood $\mathcal{S}$, with local 2-coloring $\mathcal{S}_{\mathscr{G}}$ and $\mathcal{S}_{\mathscr{R}}$, and that each B_{i-1} and each C_{i+1} is Γ-split-free and so monochromatic, and also ε-nearly in S_0 or ε-nearly outside S_0.

For each $i = 0, 1, \ldots, n-1$, let $B_{i-1}(+) \subset B_{i-1}$ be defined by

$$\mathbf{T}(B_{i-1}(+)) = A_i \cap \mathcal{S}_{\mathscr{G}},$$

so that its image under $\mathbf{T}$ is precisely the green part of A_i, and let

$$B_{i-1}(-) = B_{i-1} \setminus B_{i-1}(+).$$

Suppose that $C_{i+1} \subset \mathscr{G}_{i+1}$, so that it is on the green part of U_{i+1}. Then it follows from the definition of $\mathbf{T}$, as given by (2.1)–(2.5), that

$$C_{i+1} = \mathbf{T}^2(B_{i-1}(+)) \cup \mathbf{T}^2(B_i(-)).$$

Suppose that $C_{i+1} \subset \mathscr{R}_{i+1}$, so that it is on the red part of U_{i+1}. Corresponding to (2.4) and using the convention (2.5), let

$$i^* = \max\{j < i+1 : C_j \subset \mathscr{R}_j\}.$$

Then it follows from the definition of $\mathbf{T}$, as given by (2.1)–(2.5), that

$$C_{i^*} = \mathbf{T}^2(B_{i-1}(+)) \cup \mathbf{T}^2(B_i(-)).$$

In either case, the two sets $B_{i-1}(+)$ and $B_i(-)$, and so also the two sets B_{i-1} and B_i, are either both ε-nearly in S_0 or both ε-nearly outside S_0. Thus the sets

$$B_0, \ldots, B_{n-1} \quad \text{and} \quad C_0, \ldots, C_{n-1}$$

are either all ε-nearly in S_0 or all ε-nearly outside S_0. This contradicts the assertion in Step 1 that $\widetilde{\chi}_{S_0}(P) = k_0$ for almost every point $P \in [0,1)^2$ for some integer $k_0 < n$.

We can guarantee $\varepsilon_2 \to 0$ by taking $N \to \infty$. This establishes that $\mathbf{T} = \mathbf{T}_{\mathbf{v}}$ is ergodic, where $\mathbf{v} = (1, \mathbf{v}_0)$, for almost every $\mathbf{v}_0 \in [0,1]^2$.

Step 9. We can now use the standard technique of extending ergodicity to unique ergodicity using functional analysis and Borel measures. This completes the proof of Theorem 4.

4. Proof of Lemmas 3.2, 3.5 and 3.6

Proof of Lemma 3.2. Since S_0 is Lebesgue measurable, given any $\delta > 0$, there exists a finite set of disjoint axis-parallel rectangles such that their union V satisfies

$$\lambda_2(V \setminus S_0) + \lambda_2(S_0 \setminus V) < \delta.$$

Suppose that $0 < \lambda_2(S_0) = \tau < n$. Then

$$\lambda_2(V) > \lambda_2(S_0) - \delta = \tau - \delta \quad \text{and} \quad \lambda_2(S_0^c \cap V) < \delta,$$

where $S_0^c = X_0 \setminus S_0$. Since V is a finite union of disjoint axis-parallel rectangles, there clearly exists a threshold $t_1 = t_1(V;\delta)$ such that the union V_1 of the special $(1/t)$-squares Q contained in V has measure

$$\lambda_2(V_1) > \lambda_2(V) - \delta > \tau - 2\delta,$$

provided that the integer $t \geqslant t_1$. Let $\mathcal{B}$ denote the set of special $(1/t)$-squares Q in V_1 that satisfy

$$\frac{\lambda_2(S_0^c \cap Q)}{(1/t)^2} \geqslant \varepsilon.$$

Then, provided that $t \geqslant t_1(V;\delta)$, we have

$$\delta > \lambda_2(S_0^c \cap V) \geqslant \lambda_2(S_0^c \cap V_1) = \sum_{Q \subset V_1} \lambda_2(S_0^c \cap Q) \ \geqslant \sum_{Q \in \mathcal{B}} \lambda_2(S_0^c \cap Q) \geqslant \frac{\varepsilon|\mathcal{B}|}{t^2},$$

so that

$$|\mathcal{B}| \leqslant \frac{\delta t^2}{\varepsilon} = \frac{\varepsilon_1 t^2}{6},$$

if we choose $\delta = \varepsilon\varepsilon_1/6$. Deleting the special $(1/t)$-squares $Q \in \mathcal{B}$, we see that V_1 contains at least

$$\left(\tau - 2\delta - \frac{\varepsilon_1}{6}\right) t^2 \geqslant \left(\tau - \frac{\varepsilon_1}{2}\right) t^2 \tag{4.1}$$

special $(1/t)$-squares Q such that

$$\frac{\lambda_2(S_0^c \cap Q)}{(1/t)^2} < \varepsilon.$$

It follows that, as long as the integer $t \geqslant t_1(V;\delta)$, the number of special $(1/t)$-squares in X_0 that satisfy

$$\frac{\lambda_2(S_0 \cap Q)}{(1/t)^2} > 1 - \varepsilon \tag{4.2}$$

is bounded below by (4.1). Repeating the same argument but replacing S_0 by S_0^c, we obtain another threshold $t_2 = t_2(V;\delta)$ such that, as long as the integer $t \geqslant t_2(V;\delta)$, the number of special $(1/t)$-squares in X_0 that satisfy

$$\frac{\lambda_2(S_0 \cap Q)}{(1/t)^2} < \varepsilon \tag{4.3}$$

is bounded below by

$$\left(n - \tau - \frac{\varepsilon_1}{2}\right) t^2. \tag{4.4}$$

Combining the lower bounds (4.1) and (4.4), we see that, provided that an integer $M \geqslant \max\{t_1, t_2\}$, the number of special $(1/M)$-squares in X_0 that satisfy (4.2) or (4.3) is bounded below by $(n - \varepsilon_1)M^2$, and this completes the proof. □

Proof of Lemma 3.5. Throughout this proof, the parameters $C_1, C_2, C_3, \dots$ represent positive absolute constants.

Let $\eta = \eta(\varepsilon_2) > 0$, to be fixed later.

Suppose that an integer j^*, satisfying $0 < j^* \leqslant \eta N$, is such that both A_0 and $A_0 + j^*\mathbf{v}_0$ exhibit Γ-split, but $A_0 + j\mathbf{v}_0$ for any integer j satisfying $0 < j < j^*$ is Γ-split-free. This gives rise to a Γ-split-free chain of length j^*, and we say that $\mathbf{v}_0$ is a *bad* direction for A_0.

We now go to 3-space as follows. A discrete point sequence

$$\mathbf{s}_0 + j\mathbf{v}_0, \quad j = 0, 1, \dots, N^2 - 1,$$

in the torus $[0, 1)^2$ defines a straight line

$$\mathbf{s} + t\mathbf{v} \in \mathbb{R}^3,$$

where $\mathbf{v} = (1, \mathbf{v}_0)$. It follows that the sequence (3.3) leads to a sequence

$$A + j\mathbf{v}, \quad j = 0, 1, \dots, N^2 - 1, \tag{4.5}$$

in $\mathbb{R}^3$, where $A = \{x_0\} \times A_0$ is an axis-parallel square of side length $1/N$ on some plane $x = x_0$ where x_0 is an integer. The Γ-split-free chain under consideration is then characterized by a vector from the center of A to the centre of $A + j^*\mathbf{v}$ in $\mathbb{R}^3$. Clearly the direction of any such vector is given by $\mathbf{v}$, and the length is at most $\sqrt{3}\eta N$, where we assume that $\sqrt{3}\eta N > 1$.

Let the integer m satisfy

$$2^{m-1} < \sqrt{3}\eta N \leqslant 2^m. \tag{4.6}$$

Our basic idea is straightforward. We consider all yz-parallel unit squares arising from the lattice $\mathbb{Z}^3$, and extend Γ to each in the usual way. Assume that A is a Γ-split $(1/N)$-square on some yz-parallel unit square. Consider some other Γ-split $(1/N)$-square A^* on some other yz-parallel unit square. If the center of A^* is reached from the center of A via a vector in the direction of $\mathbf{v}$ and of length at most 2^m, then A^* determines a bad direction $\mathbf{v}$ for A. Theoretically, we can determine all possible Γ-split $(1/N)$-squares A^* on other yz-parallel unit squares, the centers of which can be reached from the center of A via vectors of length at most 2^m, leading to a collection of bad directions for A. Carrying this out, however, is impossible without more care.

Clearly every yz-parallel unit square has N^2 special $(1/N)$-squares obtained by dividing the unit square into congruent squares with side length $1/N$ in the standard way. Since the set Γ is defined by the boundaries of a finite set of polygons, there are at most $C_1 N$ special $(1/N)$-squares that exhibit Γ-split in its interior or on part of its boundary. We refer to these as Γ-split special $(1/N)$-squares. We also say that a special $(1/N)$-square on some yz-parallel unit square is *exceptional* if it has Γ-split or it is in a 3×3 array of 9 special $(1/N)$-squares, at least one of which has Γ-split. Note that some of these 9 special $(1/N)$-squares may lie on neighboring yz-parallel unit squares.

We start with a fixed Γ-split $(1/N)$-square A on some yz-parallel unit square. Assume that the bad direction vector has length between $2^{\ell-1}$ and 2^ℓ, where the positive integer $\ell \leqslant m$.

(i) Clearly $A + j^*\mathbf{v}$ is contained in a 3×3 array of 9 exceptional $(1/N)$-squares contained in at most 4 distinct but adjoining yz-parallel unit squares. Each such exceptional $(1/N)$-square contributes a set of bad directions for A with surface area measure at most

$$C_1 \left(\frac{1/N}{2^\ell} \right)^2$$

on $\mathbf{S}^2_+$.

(ii) The number of distinct yz-parallel unit squares reachable from the center of A by a vector of length between $2^{\ell-1}$ and 2^ℓ is at most $C_2(2^\ell)^3$.

(iii) The total number of exceptional $(1/N)$-squares on any yz-parallel unit square is at most C_3N.

Combining (i)–(iii), we see immediately that the total measure of bad directions for A that are characterized by bad direction vectors of length between $2^{\ell-1}$ and 2^ℓ is at most

$$C_1\left(\frac{1/N}{2^\ell}\right)^2\cdot C_2(2^\ell)^3\cdot C_3N=C_4\frac{2^\ell}{N}.$$

It follows, in view of (4.6), that the total measure of bad directions for A that are characterized by bad direction vectors of length up to ηN is at most

$$\sum_{\ell=1}^{m}C_4\frac{2^\ell}{N}<C_5\eta. \tag{4.7}$$

Next, note that A is contained in a 3×3 array of 9 exceptional $(1/N)$-squares contained in at most 4 distinct but adjoining yz-parallel unit squares. Recall that the total number of exceptional $(1/N)$-squares on any yz-parallel unit square is at most C_3N. Let $\Psi(A(\mathbf{s}_0);\mathbf{v}^*;N^2)$ denote the number of Γ-split-free chains in (4.5) of the form

$$A+j_1\mathbf{v},\ldots,A+j_2\mathbf{v},\quad 0\leqslant j_1<j_2\leqslant\eta N.$$

Then we have the inequality

$$\int_{[0,1)^2}\int_{\mathbf{S}^2}\Psi(A(\mathbf{s}_0);\mathbf{v}^*;N^2)\,\mathrm{d}\mathbf{v}^*\,\mathrm{d}\mathbf{s}_0\leqslant C_3N\cdot C_5\eta=C_6\eta N. \tag{4.8}$$

Remark. An inequality of the form

$$\int_{\mathbf{S}^2}\Psi(A(\mathbf{s}_0);\mathbf{v}^*;N^2)\,\mathrm{d}\mathbf{v}^*\leqslant C_5\eta$$

for every $A(\mathbf{s}^*)$ would be ideal, but cannot be deduced from the upper bound (4.7), in view of the possibility that the integrand may exceed 1. Thus we need to average over $\mathbf{s}_0$ as well. The inequality (4.8) is valid with the integrand $\Psi(A(\mathbf{s}^*);\mathbf{v}^*;N^2)$, since if the sequence has more than one Γ-split-free chain, the multiplicity is taken care of by the parameter $\mathbf{s}_0$, as the terms of the sequence (4.5) correspond to N^2 distinct values of $\mathbf{s}_0$.

Let $\Phi(\varepsilon_2N;\eta N)$ denote the set of pairs $(\mathbf{s}_0,\mathbf{v}^*)\in[0,1)^2\times\mathbf{S}^2_+$ such that the sequence (4.5) has more than ε_2N Γ-split-free chains with length at most ηN. Then writing meas for the product measure, we have

$$\operatorname{meas}(\Phi(\varepsilon_2N;\eta N))\varepsilon_2N\leqslant\int_{[0,1)^2}\int_{\mathbf{S}^2}\Psi(A(\mathbf{s}_0);\mathbf{v}^*;N^2)\,\mathrm{d}\mathbf{v}^*\,\mathrm{d}\mathbf{s}_0. \tag{4.9}$$

Combining (4.8) and (4.9), we conclude that

$$\operatorname{meas}(\Phi(\varepsilon_2N;\eta N))\leqslant\frac{C_6\eta}{\varepsilon_2}.$$

The proof is now complete if we choose $\eta=c_3\varepsilon_2^2$ with a suitable constant c_3. □

Proof of Lemma 3.6. For any integer $k=1,\ldots,N^2$, write

$$\Omega(N;k)=\left\{\mathbf{v}_0=(\alpha_1,\alpha_2)\in[0,1]^2:\|k\alpha_1\|<\frac{2}{N}\text{ and }\|k\alpha_2\|<\frac{2}{N}\right\}. \tag{4.10}$$

It is not difficult to see that

$$\Omega(N;k)\cap[0,1)^2=\left(\left[0,\frac{2}{kN}\right)\cup\bigcup_{j=1}^{k-1}\left(\frac{j}{k}-\frac{2}{kN},\frac{j}{k}+\frac{2}{kN}\right)\cup\left(1-\frac{2}{kN},1\right)\right)^2,$$

so that

$$\lambda_2(\Omega(N;k))=\left(\frac{4}{N}\right)^2=\frac{16}{N^2}. \tag{4.11}$$

For every $(\alpha_1,\alpha_2)\in[0,1]^2$, consider the counting function

$$\omega_N(\alpha_1,\alpha_2)=|\{k=1,\ldots,N^2:(\alpha_1,\alpha_2)\in\Omega(N;k)\}|. \tag{4.12}$$

Combining (4.10)–(4.12), we see that

$$\int_0^1\int_0^1\omega_N(\alpha_1,\alpha_2)\,\mathrm{d}\alpha_1\,\mathrm{d}\alpha_2=\sum_{k=1}^{N^2}\lambda_2(\Omega(N;k))=16.$$

Given any $\varepsilon_2>0$, write

$$\Omega(\varepsilon_2)=\left\{(\alpha_1,\alpha_2)\in[0,1]^2:\omega_N(\alpha_1,\alpha_2)\geqslant\frac{16}{\varepsilon_2}\right\}.$$

Since the function $\omega_N(\alpha_1,\alpha_2)$ is non-negative, we clearly have

$$16=\int_0^1\int_0^1\omega_N(\alpha_1,\alpha_2)\,\mathrm{d}\alpha_1\,\mathrm{d}\alpha_2\geqslant\frac{16}{\varepsilon_2}\lambda_2(\Omega(\varepsilon_2)),$$

so that $\lambda_2(\Omega(\varepsilon_2))\leqslant\varepsilon_2$. It follows that

$$\lambda_2([0,1]^2\setminus\Omega(\varepsilon_2))>1-\varepsilon_2,$$

so that the collection of vectors

$$\mathbf{v}_0=(\alpha_1,\alpha_2)\in[0,1]^2\setminus\Omega(\varepsilon_2) \tag{4.13}$$

represents at least $(1-\varepsilon_2)$-proportion of the set $[0,1]^2$.

We shall show that the lemma holds with the choice

$$C^*=C^*(\varepsilon_2)=1+\frac{16}{\varepsilon_2}.$$

Suppose on the contrary that some axis-parallel square Q with side length $2/N$ contains modulo one more than $1+16/\varepsilon_2$ elements of some sequence

$$\mathbf{s}_0+j\mathbf{v}_0,\quad j=0,1,\ldots,N^2-1.$$

In other words, suppose that there exists a subset $J\subset\{0,1,\ldots,N^2-1\}$ with $|J|>1+16/\varepsilon_2$ such that

$$\mathbf{s}_0+j\mathbf{v}_0\in Q,\quad j\in J.$$

Let j_0 be the smallest element of J, and let

$$J^*=J\setminus\{j_0\}.$$

It is not difficult to see that

$$J^*\subset\{k=1,\ldots,N^2:(\alpha_1,\alpha_2)\in\Omega(N;k)\},$$

and so

$$\omega_N(\mathbf{v}_0)\geqslant|J^*|>\frac{16}{\varepsilon_2}.$$

This implies that $\mathbf{v}_0\in\Omega(\varepsilon_2)$, contradicting (4.13). The lemma now follows. $\square$

5. Establishing Lemma 3.4

In this section, we establish an intermediate result from which Lemma 3.4 follows easily.

Lemma 5.1. *Let the integer H be even and positive, and let $\Delta = \Delta(u_1, u_2; \theta)$ be an arbitrary rectangle with side lengths $2u_1$ and $2u_2$, and tilted by an angle θ in the anticlockwise direction, where $0 < u_1, u_2 < 1/2$. Let*

$$F(\mathbf{s}_0; \alpha_1, \alpha_2; H) = |\{j = 0, 1, \ldots, H-1 : \mathbf{s}_0 + j(\alpha_1, \alpha_2) \in \Delta\}|.$$

Then for any parameter $\kappa > 1$, we have

$$\lambda_2\left(\left\{(\alpha_1, \alpha_2) \in [0,1]^2 : F(\mathbf{s}_0; \alpha_1, \alpha_2; H) \geqslant \frac{u_1 u_2 H}{2} - \kappa\psi(u_1, u_2)\right\}\right) \geqslant 1 - \frac{1}{\kappa},$$

where

$$\psi(u_1, u_2) = 2^{23} \max\left\{\frac{u_1}{u_2^{1/2}}, \frac{u_2}{u_1^{1/2}}\right\} + \frac{2^{16}}{u_1^{1/3} u_2^{1/3}}. \tag{5.1}$$

Proof of Lemma 3.4. We take $H = N^2$. For the buffer zone $\Delta = \mathcal{B}_N$, we have

$$u_1 = \frac{c_1}{2}, \quad \text{and} \quad u_2 = \frac{1}{4N}.$$

Then

$$\frac{u_1 u_2 H}{2} = \frac{c_1 N}{16} \quad \text{and} \quad \psi(u_1, u_2) = c_4 N^{1/2} + c_5 N^{1/3} \tag{5.2}$$

for some positive absolute constants c_4 and c_5. For large values of N, the right hand side in (5.2) is much smaller than the left hand side. Thus given any $\varepsilon_2 > 0$, we can choose a sufficiently large $\kappa = \kappa(\varepsilon_2) > 1$ to guarantee that the inequality (3.2) holds for some constant c_2 satisfying $0 < c_2 < c_1/16$. □

Proof of Lemma 5.1. We proceed by a number of steps.

Step 1. We aim to give a good description of the term $F(\mathbf{s}_0; \alpha_1, \alpha_2; H)$.

Let $\mathcal{L}(\alpha_1, \alpha_2)$ denote the lattice in $\mathbb{R}^3$ generated by the vectors

$$\mathbf{e}_1 = (\alpha_1, \alpha_2, 1), \quad \mathbf{e}_2 = (-1, 0, 0), \quad \mathbf{e}_3 = (0, -1, 0),$$

and consider the 3×3 matrix

$$M = \begin{pmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \end{pmatrix} = \begin{pmatrix} \alpha_1 & -1 & 0 \\ \alpha_2 & 0 & -1 \\ 1 & 0 & 0 \end{pmatrix}, \quad \text{with } \det(M) = 1. \tag{5.3}$$

Then, with $\mathbf{n} \in \mathbb{Z}^3$ expressed as column vectors, we have

$$\mathcal{L}(\alpha_1, \alpha_2) = \{M\mathbf{n} : \mathbf{n} \in \mathbb{Z}^3\}. \tag{5.4}$$

Writing $B = (\Delta - \mathbf{s}_0) \times [0, H) \subset \mathbb{R}^3$, we then have

$$F(\mathbf{s}_0; \alpha_1, \alpha_2; H) = |\mathcal{L}(\alpha_1, \alpha_2) \cap B| = \sum_{\mathbf{n} \in \mathbb{Z}^3} \chi_B(M\mathbf{n}), \tag{5.5}$$

where χ_B is the characteristic function of B.

We next use the Poisson summation formula, that under some mild condition on a function $f : \mathbb{R}^3 \to \mathbb{R}$, we have

$$\sum_{\mathbf{n} \in \mathbb{Z}^3} f(\mathbf{n}) = \sum_{\mathbf{m} \in \mathbb{Z}^3} \int_{\mathbb{R}^3} f(\mathbf{x}) \mathrm{e}^{-2\pi \mathrm{i} \mathbf{x} \cdot \mathbf{m}} \, \mathrm{d}\mathbf{x}. \tag{5.6}$$

Applying this formula with $f = \chi_B$, it follows from (5.5) and noting (5.3) that

$$\begin{aligned} F(\mathbf{s}_0; \alpha_1, \alpha_2; H) &= \sum_{\mathbf{n}\in\mathbb{Z}^3} \chi_B(M\mathbf{n}) = \sum_{\mathbf{m}\in\mathbb{Z}^3} \int_{\mathbb{R}^3} \chi_B(M\mathbf{x}) \mathrm{e}^{-2\pi\mathrm{i}\mathbf{x}\cdot\mathbf{m}}\,\mathrm{d}\mathbf{x} \\ &= \frac{1}{\det(M)} \sum_{\mathbf{m}\in\mathbb{Z}^3} \int_{\mathbb{R}^3} \chi_B(\mathbf{z}) \mathrm{e}^{-2\pi\mathrm{i}M^{-1}\mathbf{z}\cdot\mathbf{m}}\,\mathrm{d}\mathbf{z} = \sum_{\mathbf{m}\in\mathbb{Z}^3} \int_B \mathrm{e}^{-2\pi\mathrm{i}M^{-1}\mathbf{z}\cdot\mathbf{m}}\,\mathrm{d}\mathbf{z}. \end{aligned} \tag{5.7}$$

Note, in particular, that for $\mathbf{m} = \mathbf{0}$, we have

$$\int_B \mathrm{e}^{-2\pi\mathrm{i}M^{-1}\mathbf{z}\cdot\mathbf{m}}\,\mathrm{d}\mathbf{z} = \int_B \mathrm{e}^{-2\pi\mathrm{i}M^{-1}\mathbf{z}\cdot\mathbf{0}}\,\mathrm{d}\mathbf{z} = \lambda_3(B) = \lambda_2(\Delta)H = 4u_1u_2H. \tag{5.8}$$

On the other hand, it is easy to see that the inverse matrix

$$M^{-1} = \begin{pmatrix} 0 & 0 & 1 \\ -1 & 0 & \alpha_1 \\ 0 & -1 & \alpha_2 \end{pmatrix},$$

so that $M^{-1}\mathbf{z} = (z_3, \alpha_1 z_3 - z_1, \alpha_2 z_3 - z_2)$. Thus with $\mathbf{v} = (1, \alpha_1, \alpha_2)$, we have

$$\begin{aligned} M^{-1}\mathbf{z}\cdot\mathbf{m} &= z_3 m_1 + (\alpha_1 z_3 - z_1)m_2 + (\alpha_2 z_3 - z_2)m_3 \\ &= -z_1 m_2 - z_2 m_3 + z_3(m_1 + \alpha_1 m_2 + \alpha_2 m_3) \\ &= -z_1 m_2 - z_2 m_3 + z_3 \mathbf{v}\cdot\mathbf{m}. \end{aligned} \tag{5.9}$$

Then for every $\mathbf{m} \in \mathbb{Z}^3 \setminus \{\mathbf{0}\}$, we can write

$$\int_B \mathrm{e}^{-2\pi\mathrm{i}M^{-1}\mathbf{z}\cdot\mathbf{m}}\,\mathrm{d}\mathbf{z} = \mathcal{I}(\mathbf{m}; \Delta - \mathbf{s}_0)\mathcal{J}(\mathbf{m}; \mathbf{v}; H), \tag{5.10}$$

where

$$\mathcal{I}(\mathbf{m}; \Delta - \mathbf{s}_0) = \int_{\Delta-\mathbf{s}_0} \mathrm{e}^{2\pi\mathrm{i}(z_1 m_2 + z_2 m_3)}\,\mathrm{d}z_1\,\mathrm{d}z_2 \tag{5.11}$$

and

$$\mathcal{J}(\mathbf{m}; \mathbf{v}; H) = \int_0^H \mathrm{e}^{-2\pi\mathrm{i}z_3\mathbf{v}\cdot\mathbf{m}}\,\mathrm{d}z_3 = \frac{1 - \mathrm{e}^{-2\pi\mathrm{i}H\mathbf{v}\cdot\mathbf{m}}}{2\pi\mathrm{i}\mathbf{v}\cdot\mathbf{m}}. \tag{5.12}$$

Write

$$\mathbf{z}^* = (z_1, z_2) \quad \text{and} \quad \mathbf{m}^* = (m_2, m_3),$$

and let

$$\Delta = \rho([-u_1, u_1] \times [-u_2, u_2]) + \mathbf{w} \subset [0, 1)^2,$$

where ρ denotes an anticlockwise rotation. Then

$$\mathbf{z}^* \in \Delta - \mathbf{s}_0 = \rho([-u_1, u_1] \times [-u_2, u_2]) + \mathbf{w} - \mathbf{s}_0$$

if and only if

$$\mathbf{z}' = (z_1', z_2') = \rho^{-1}(\mathbf{z}^* - \mathbf{w} + \mathbf{s}_0) \in [-u_1, u_1] \times [-u_2, u_2].$$

Note that

$$\mathbf{z}' = \rho^{-1}(\mathbf{z}^* - \mathbf{w} + \mathbf{s}_0) \quad \text{if and only if} \quad \mathbf{z}^* = \rho\mathbf{z}' + \mathbf{w} - \mathbf{s}_0.$$

It then follows from (5.11) that

$$\begin{aligned}\mathcal{I}(\mathbf{m};\Delta-\mathbf{s}_0)&=\int_{\Delta-\mathbf{s}_0}\mathrm{e}^{2\pi\mathrm{i}\mathbf{z}^*\cdot\mathbf{m}^*}\,\mathrm{d}\mathbf{z}^*=\int_{[-u_1,u_1]\times[-u_2,u_2]}\mathrm{e}^{2\pi\mathrm{i}(\rho\mathbf{z}'+\mathbf{w}-\mathbf{s}_0)\cdot\mathbf{m}^*}\,\mathrm{d}\mathbf{z}'\\&=\mathrm{e}^{2\pi\mathrm{i}(\mathbf{w}-\mathbf{s}_0)\cdot\mathbf{m}^*}\int_{[-u_1,u_1]\times[-u_2,u_2]}\mathrm{e}^{2\pi\mathrm{i}\rho\mathbf{z}'\cdot\mathbf{m}^*}\,\mathrm{d}\mathbf{z}'\\&=\mathrm{e}^{2\pi\mathrm{i}(\mathbf{w}-\mathbf{s}_0)\cdot\mathbf{m}^*}\int_{[-u_1,u_1]\times[-u_2,u_2]}\mathrm{e}^{2\pi\mathrm{i}\mathbf{z}'\cdot\rho^{-1}\mathbf{m}^*}\,\mathrm{d}\mathbf{z}',\end{aligned}\tag{5.13}$$

since $\rho\mathbf{z}'\cdot\mathbf{m}^*=\mathbf{z}'\cdot\rho^{-1}\mathbf{m}^*$ as the scalar product remains unchanged under identical rotation for both constituent vectors. Furthermore, if ρ is anticlockwise rotation by an angle θ, then

$$\rho^{-1}\mathbf{m}^*=\rho^{-1}(m_2,m_3)=(m_2\cos\theta+m_3\sin\theta,m_3\cos\theta-m_2\sin\theta).\tag{5.14}$$

Combining (5.13) and (5.14), we conclude that

$$\mathcal{I}(\mathbf{m};\Delta-\mathbf{s}_0)=\mathrm{e}^{2\pi\mathrm{i}(\mathbf{w}-\mathbf{s}_0)\cdot\mathbf{m}^*}\mathcal{I}(m_2,m_3;\theta;u_1)\mathcal{I}(m_3,-m_2;\theta;u_2),\tag{5.15}$$

where

$$\begin{aligned}\mathcal{I}(m_2,m_3;\theta;u_1)&=\int_{-u_1}^{u_1}\mathrm{e}^{2\pi\mathrm{i}z_1'(m_2\cos\theta+m_3\sin\theta)}\,\mathrm{d}z_1'\\&=2\int_0^{u_1}\cos(2\pi z_1'(m_2\cos\theta+m_3\sin\theta))\,\mathrm{d}z_1'\\&=\frac{\sin(2\pi u_1(m_2\cos\theta+m_3\sin\theta))}{\pi(m_2\cos\theta+m_3\sin\theta)}\end{aligned}\tag{5.16}$$

and

$$\begin{aligned}\mathcal{I}(m_3,-m_2;\theta;u_2)&=\int_{-u_2}^{u_2}\mathrm{e}^{2\pi\mathrm{i}z_2'(m_3\cos\theta-m_2\sin\theta)}\,\mathrm{d}z_2'\\&=2\int_0^{u_2}\cos(2\pi z_2'(m_3\cos\theta-m_2\sin\theta))\,\mathrm{d}z_2'\\&=\frac{\sin(2\pi u_2(m_3\cos\theta-m_2\sin\theta))}{\pi(m_3\cos\theta-m_2\sin\theta)}.\end{aligned}\tag{5.17}$$

Finally, combining (5.7), (5.8), (5.10) and (5.15), we conclude that

$$\begin{aligned}&F(\mathbf{s}_0;\alpha_1,\alpha_2;H)-4u_1u_2H\\&\quad=\sum_{\mathbf{m}\in\mathbb{Z}^3\setminus\{\mathbf{0}\}}\mathrm{e}^{2\pi\mathrm{i}(\mathbf{w}-\mathbf{s}_0)\cdot\mathbf{m}^*}\mathcal{I}(m_2,m_3;\theta;u_1)\mathcal{I}(m_3,-m_2;\theta;u_2)\mathcal{J}(\mathbf{m};\mathbf{v};H),\end{aligned}\tag{5.18}$$

where the details for the various factors in the summand are given by (5.12), (5.16) and (5.17).

Step 2. Here we contract the interval $[0,H]$ in the third direction and average over all contractions. More precisely, for every $h\in\mathbb{R}$ satisfying $0\leqslant h\leqslant H/2$, consider the smaller set

$$B(h)=(\Delta-\mathbf{s}_0)\times[h,H-h)\subset B.$$

Clearly, in view of an analog of (5.8), we have

$$
\begin{aligned}
F(\mathbf{s}_0;\alpha_1,\alpha_2;H) &\geqslant \frac{2}{H}\int_0^{H/2}\left(\sum_{\mathbf{n}\in\mathbb{Z}^3}\chi_{B(h)}(M\mathbf{n})\right)\mathrm{d}h \\
&= \frac{2}{H}\int_0^{H/2} 4u_1u_2(H-2h)\,\mathrm{d}h + \frac{2}{H}\int_0^{H/2}\left(\sum_{\mathbf{m}\in\mathbb{Z}^3\setminus\{\mathbf{0}\}}\int_{B(h)}\mathrm{e}^{-2\pi\mathrm{i}M^{-1}\mathbf{z}\cdot\mathbf{m}}\,\mathrm{d}\mathbf{z}\right)\mathrm{d}h \\
&= 2u_1u_2H + \frac{2}{H}\int_0^{H/2}\left(\sum_{\mathbf{m}\in\mathbb{Z}^3\setminus\{\mathbf{0}\}}\int_{B(h)}\mathrm{e}^{-2\pi\mathrm{i}M^{-1}\mathbf{z}\cdot\mathbf{m}}\,\mathrm{d}\mathbf{z}\right)\mathrm{d}h. \qquad (5.19)
\end{aligned}
$$

For any $\mathbf{m}\in\mathbb{Z}^3\setminus\{\mathbf{0}\}$, analogous to (5.10), we have

$$
\int_{B(h)}\mathrm{e}^{-2\pi\mathrm{i}M^{-1}\mathbf{z}\cdot\mathbf{m}}\,\mathrm{d}\mathbf{z} = \mathcal{I}(\mathbf{m};\Delta-\mathbf{s}_0)\mathcal{J}(\mathbf{m};\mathbf{v};H;h) \qquad (5.20)
$$

where the first two directions are unaffected and

$$
\mathcal{J}(\mathbf{m};\mathbf{v};H;h) = \int_h^{H-h}\mathrm{e}^{-2\pi\mathrm{i}z_3\mathbf{v}\cdot\mathbf{m}}\,\mathrm{d}z_3 = \frac{\mathrm{e}^{-2\pi\mathrm{i}h\mathbf{v}\cdot\mathbf{m}}-\mathrm{e}^{-2\pi\mathrm{i}(H-h)\mathbf{v}\cdot\mathbf{m}}}{2\pi\mathrm{i}\mathbf{v}\cdot\mathbf{m}}, \qquad (5.21)
$$

so that

$$
\int_0^{H/2}\mathcal{J}(\mathbf{m};\mathbf{v};H;h)\,\mathrm{d}h = \frac{1-2\mathrm{e}^{-\pi\mathrm{i}H\mathbf{v}\cdot\mathbf{m}}+\mathrm{e}^{-2\pi\mathrm{i}H\mathbf{v}\cdot\mathbf{m}}}{(2\pi\mathrm{i}\mathbf{v}\cdot\mathbf{m})^2} = -\frac{(1-\mathrm{e}^{-\pi\mathrm{i}H\mathbf{v}\cdot\mathbf{m}})^2}{4\pi^2(\mathbf{v}\cdot\mathbf{m})^2},
$$

and so

$$
\widetilde{\mathcal{J}}(\mathbf{m};\mathbf{v};H) = \frac{2}{H}\int_0^{H/2}\mathcal{J}(\mathbf{m};\mathbf{v};H;h)\,\mathrm{d}h = -\frac{(1-\mathrm{e}^{-\pi\mathrm{i}H\mathbf{v}\cdot\mathbf{m}})^2}{2\pi^2H(\mathbf{v}\cdot\mathbf{m})^2}. \qquad (5.22)
$$

Finally, combining (5.15), (5.19), (5.20) and (5.22), we conclude that

$$
\begin{aligned}
&F(\mathbf{s}_0;\alpha_1,\alpha_2;H) - 2u_1u_2H \\
&\quad\geqslant \sum_{\mathbf{m}\in\mathbb{Z}^3\setminus\{\mathbf{0}\}}\mathrm{e}^{2\pi\mathrm{i}(\mathbf{w}-\mathbf{s}_0)\cdot\mathbf{m}^*}\mathcal{I}(m_2,m_3;\theta;u_1)\mathcal{I}(m_3,-m_2;\theta;u_2)\widetilde{\mathcal{J}}(\mathbf{m};\mathbf{v};H), \qquad (5.23)
\end{aligned}
$$

the analog of (5.18).

Step 3. Here we also contract the rectangle $[-u_1,u_1]\times[-u_2,u_2]$ in the first two directions and average over all contractions. More precisely, for every γ_1 and γ_2 satisfying $0\leqslant\gamma_1,\gamma_2\leqslant 1$, consider the smaller set

$$
\Delta(\gamma_1,\gamma_2) = \rho([-\gamma_1u_1,\gamma_1u_1]\times[-\gamma_2u_2,\gamma_2u_2]) + \mathbf{w}\subset[0,1)^2,
$$

and the corresponding smaller set

$$
B(\gamma_1,\gamma_2;h) = (\Delta(\gamma_1,\gamma_2)-\mathbf{s}_0)\times[h,H-h)\subset B(h)\subset B.
$$

Clearly, in view of an analog of (5.8), we have

$$
\begin{aligned}
F(\mathbf{s}_0;\alpha_1,\alpha_2;H) &\geqslant \frac{2}{H}\int_0^{H/2}\int_0^1\int_0^1\left(\sum_{\mathbf{n}\in\mathbb{Z}^3}\chi_{B(\gamma_1,\gamma_2,h)}(M\mathbf{n})\right)\mathrm{d}\gamma_1\,\mathrm{d}\gamma_2\,\mathrm{d}h \\
&= \frac{2}{H}\int_0^{H/2}\int_0^1\int_0^1 4\gamma_1\gamma_2u_1u_2(H-2h)\,\mathrm{d}\gamma_1\,\mathrm{d}\gamma_2\,\mathrm{d}h \\
&\quad + \frac{2}{H}\int_0^{H/2}\int_0^1\int_0^1\left(\sum_{\mathbf{m}\in\mathbb{Z}^3\setminus\{\mathbf{0}\}}\int_{B(\gamma_1,\gamma_2,h)}\mathrm{e}^{-2\pi\mathrm{i}M^{-1}\mathbf{z}\cdot\mathbf{m}}\,\mathrm{d}\mathbf{z}\right)\mathrm{d}\gamma_1\,\mathrm{d}\gamma_2\,\mathrm{d}h, \qquad (5.24)
\end{aligned}
$$

where the first term

$$\frac{2}{H}\int_0^{H/2}\int_0^1\int_0^1 4\gamma_1\gamma_2u_1u_2(H-2h)\,\mathrm{d}\gamma_1\,\mathrm{d}\gamma_2\,\mathrm{d}h=\frac{u_1u_2H}{2}. \tag{5.25}$$

For any $\mathbf{m}\in\mathbb{Z}^3\setminus\{\mathbf{0}\}$, analogous to (5.10) and (5.20), we have

$$\int_{B(\gamma_1,\gamma_2,h)}\mathrm{e}^{-2\pi\mathrm{i}M^{-1}\mathbf{z}\cdot\mathbf{m}}\,\mathrm{d}\mathbf{z}=\mathcal{I}(\mathbf{m};\Delta(\gamma_1,\gamma_2)-\mathbf{s}_0)\mathcal{J}(\mathbf{m};\mathbf{v};H;h) \tag{5.26}$$

Then, analogous to (5.15), we have

$$\mathcal{I}(\mathbf{m};\Delta(\gamma_1,\gamma_2)-\mathbf{s}_0)=\mathrm{e}^{2\pi\mathrm{i}(\mathbf{w}-\mathbf{s}_0)\cdot\mathbf{m}^*}\mathcal{I}(m_2,m_3;\theta;\gamma_1u_1)\mathcal{I}(m_3,-m_2;\theta;\gamma_2u_2). \tag{5.27}$$

Simple calculations now give

$$\begin{aligned}\widetilde{\mathcal{I}}(m_2,m_3;\theta;u_1)&=\int_0^1\mathcal{I}(m_2,m_3;\theta;\gamma_1u_1)\,\mathrm{d}\gamma_1\\&=\int_0^1\frac{\sin(2\pi\gamma_1u_1(m_2\cos\theta+m_3\sin\theta))}{\pi(m_2\cos\theta+m_3\sin\theta)}\,\mathrm{d}\gamma_1\\&=\frac{1-\cos(2\pi u_1(m_2\cos\theta+m_3\sin\theta))}{2\pi^2u_1(m_2\cos\theta+m_3\sin\theta)^2}\\&=\frac{\sin^2(\pi u_1(m_2\cos\theta+m_3\sin\theta))}{\pi^2u_1(m_2\cos\theta+m_3\sin\theta)^2}\end{aligned} \tag{5.28}$$

and

$$\begin{aligned}\widetilde{\mathcal{I}}(m_3,-m_2;\theta;u_2)&=\int_0^1\mathcal{I}(m_3,-m_2;\theta;\gamma_2u_2)\,\mathrm{d}\gamma_2\\&=\int_0^1\frac{\sin(2\pi\gamma_2u_2(m_3\cos\theta-m_2\sin\theta))}{\pi(m_3\cos\theta-m_2\sin\theta)}\,\mathrm{d}\gamma_2\\&=\frac{1-\cos(2\pi u_2(m_3\cos\theta-m_2\sin\theta))}{2\pi^2u_2(m_3\cos\theta-m_2\sin\theta)^2}\\&=\frac{\sin^2(\pi u_2(m_3\cos\theta-m_2\sin\theta))}{\pi^2u_2(m_3\cos\theta-m_2\sin\theta)^2}.\end{aligned} \tag{5.29}$$

It now follows from (5.22) and (5.24)–(5.29) that

$$\begin{aligned}&F(\mathbf{s}_0;\alpha_1,\alpha_2;H)-\frac{u_1u_2H}{2}\\&\quad\geqslant\sum_{\mathbf{m}\in\mathbb{Z}^3\setminus\{\mathbf{0}\}}\mathrm{e}^{2\pi\mathrm{i}(\mathbf{w}-\mathbf{s}_0)\cdot\mathbf{m}^*}\widetilde{\mathcal{I}}(m_2,m_3;\theta;u_1)\widetilde{\mathcal{I}}(m_3,-m_2;\theta;u_2)\widetilde{\mathcal{J}}(\mathbf{m};\mathbf{v};H),\end{aligned} \tag{5.30}$$

the analog of (5.18) and (5.23).

Step 4. It follows from (5.30) that

$$F(\mathbf{s}_0;\alpha_1,\alpha_2;H)\geqslant\frac{u_1u_2H}{2}+\sum_{\mathbf{m}\in\mathbb{Z}^3\setminus\{\mathbf{0}\}}\Lambda(\theta;u_1,u_2;H;\mathbf{v};\mathbf{m}),$$

where

$$\Lambda(\theta;u_1,u_2;H;\mathbf{v};\mathbf{m})=\widetilde{\mathcal{I}}(m_2,m_3;\theta;u_1)\widetilde{\mathcal{I}}(m_3,-m_2;\theta;u_2)|\widetilde{\mathcal{J}}(\mathbf{m};\mathbf{v};H)|, \tag{5.31}$$

since it is clear from (5.28) and (5.29) that $\widetilde{\mathcal{I}}(m_2,m_3;\theta;u_1)$ and $\widetilde{\mathcal{I}}(m_3,-m_2;\theta;u_2)$ are real and non-negative.

The inequalities in Steps 1–3 are at this stage only formal inequalities, as we have not considered the question of convergence. Write

$$\Xi(\alpha_1,\alpha_2)=\sum_{\mathbf{m}\in\mathbb{Z}^3\setminus\{\mathbf{0}\}}\Lambda(\theta;u_1,u_2;H;\mathbf{v};\mathbf{m}), \tag{5.32}$$

where $\mathbf{v}=(1,\alpha_1,\alpha_2)$. We shall use the first-moment method and analyze the average

$$\int_0^1\int_0^1\Xi(\alpha_1,\alpha_2)\,\mathrm{d}\alpha_1\,\mathrm{d}\alpha_2,$$

and remove those directions (α_1,α_2) for which $\Xi(\alpha_1,\alpha_2)$ is substantially larger than the average. In this step, we do some further preparation.

First of all, the inequalities

$$\widetilde{\mathcal{I}}(m_2,m_3;\theta;u_1)\leqslant\min\left\{u_1,\frac{1}{\pi^2u_1(m_2\cos\theta+m_3\sin\theta)^2}\right\} \tag{5.33}$$

and

$$\widetilde{\mathcal{I}}(m_3,-m_2;\theta;u_2)\leqslant\min\left\{u_2,\frac{1}{\pi^2u_2(m_3\cos\theta-m_2\sin\theta)^2}\right\} \tag{5.34}$$

follow on applying the inequality $|\sin y|\leqslant\min\{|y|,1\}$, which holds for every $y\in\mathbb{R}$, to (5.28) and (5.29) respectively. Next, the inequality

$$|\widetilde{\mathcal{J}}(\mathbf{m};\mathbf{v};H)|=\frac{|1-\mathrm{e}^{-\pi\mathrm{i}H\mathbf{v}\cdot\mathbf{m}}|^2}{2\pi^2H(\mathbf{v}\cdot\mathbf{m})^2}\leqslant\min\left\{\frac{H}{2},\frac{2}{\pi^2H(\mathbf{v}\cdot\mathbf{m})^2}\right\} \tag{5.35}$$

follows on applying the inequality $|1-\mathrm{e}^{\mathrm{i}y}|\leqslant\min\{|y|,2\}$, which also holds for every $y\in\mathbb{R}$, to (5.22). Since we consider only those $\mathbf{m}\neq\mathbf{0}$, it follows that in the trivial case when $\mathbf{m}^*=(m_2,m_3)=(0,0)$, we must have $m_1\neq0$ and, since H is even, also

$$|\widetilde{\mathcal{J}}(\mathbf{m};\mathbf{v};H)|=\frac{|1-\mathrm{e}^{-\pi\mathrm{i}H\mathbf{v}\cdot\mathbf{m}}|^2}{2\pi^2H(\mathbf{v}\cdot\mathbf{m})^2}=0. \tag{5.36}$$

Hence we can assume that $\mathbf{m}^*=(m_2,m_3)\in\mathbb{Z}^2\setminus\{(0,0)\}$. We can write

$$\mathbb{Z}^2\setminus\{(0,0)\}=\bigcup_{j=0}^{\infty}\mathcal{Z}_j \tag{5.37}$$

as a disjoint union of subsets, where for every $j=0,1,2,3,\ldots$, the subset

$$\begin{aligned}\mathcal{Z}_j&=\{\mathbf{m}^*=(m_2,m_3)\in\mathbb{Z}^2\setminus\{(0,0)\}:2^{j-1}<\max\{|m_2|,|m_3|\}\leqslant2^j\}\\&=\mathbb{Z}^2\cap([-2^j,2^j]^2\setminus[-2^{j-1},2^{j-1}]^2).\end{aligned} \tag{5.38}$$

For every $\mathbf{m}^*=(m_2,m_3)\in\mathcal{Z}_j$, let

$$\Omega_j(\mathbf{m}^*;0)=\left\{(\alpha_1,\alpha_2)\in[0,1]^2:|\mathbf{v}\cdot\mathbf{m}|\leqslant\frac{1}{H}\text{ for some }m_1\in\mathbb{Z}\right\}, \tag{5.39}$$

and for every integer $\ell\geqslant1$, let

$$\Omega_j(\mathbf{m}^*;\ell)=\left\{(\alpha_1,\alpha_2)\in[0,1]^2:\frac{2^{\ell-1}}{H}<|\mathbf{v}\cdot\mathbf{m}|\leqslant\frac{2^\ell}{H}\text{ for some }m_1\in\mathbb{Z}\right\}. \tag{5.40}$$

We claim that the inequality

$$\lambda_2(\Omega_j(\mathbf{m}^*;\ell))\leqslant\min\left\{\frac{2^{\ell+5}}{H}+\frac{2^{2\ell+3-j}}{H^2},1\right\}\leqslant\min\left\{\frac{2^{\ell+6}}{H},1\right\} \tag{5.41}$$

holds for every $\mathbf{m}^*=(m_2,m_3)\in\mathcal{Z}_j$ and every integer $\ell\geqslant0$.

Suppose first of all that $\ell = 0$. To estimate $\lambda_2(\Omega_j(\mathbf{m}^*;0))$, we may assume that $|m_2| \leqslant |m_3|$. Then for fixed m_1, α_1, m_2, the variable $\alpha_2 \in [0,1]$ must satisfy the inequality

$$-\frac{1}{H} \leqslant m_1 + \alpha_1 m_2 + \alpha_2 m_3 \leqslant \frac{1}{H}, \tag{5.42}$$

and so falls into an interval of length

$$\frac{2}{Hm_3} \leqslant \frac{2^{2-j}}{H}.$$

Note next that the inequality (5.42) has no solution if

$$|m_1| > 2^{j+1} + \frac{1}{H}.$$

It follows that

$$\lambda_2(\Omega_j(\mathbf{m}^*;0)) \leqslant \min\left\{\left(2^{j+2} + 1 + \frac{2}{H}\right)\frac{2^{2-j}}{H}, 1\right\} \leqslant \min\left\{\frac{2^5}{H} + \frac{2^{3-j}}{H^2}, 1\right\},$$

and this establishes (5.41) for $\ell = 0$.

Suppose next that $\ell \geqslant 1$. To estimate $\lambda_2(\Omega_j(\mathbf{m}^*;\ell))$, we may assume again that $|m_2| \leqslant |m_3|$. Then for fixed m_1, α_1, m_2, the variable $\alpha_2 \in [0,1]$ must satisfy one of the inequalities

$$-\frac{2^\ell}{H} \leqslant m_1 + \alpha_1 m_2 + \alpha_2 m_3 < -\frac{2^{\ell-1}}{H} \quad \text{or} \quad \frac{2^{\ell-1}}{H} < m_1 + \alpha_1 m_2 + \alpha_2 m_3 \leqslant \frac{2^\ell}{H}, \tag{5.43}$$

and so falls into two intervals of total length

$$\frac{2^\ell}{Hm_3} \leqslant \frac{2^{\ell+1-j}}{H}.$$

Note next that the inequalities in (5.43) have no solution if

$$|m_1| > 2^{j+1} + \frac{2^\ell}{H}.$$

It follows that

$$\lambda_2(\Omega_j(\mathbf{m}^*;\ell)) \leqslant \min\left\{\left(2^{j+2} + 1 + \frac{2^{\ell+1}}{H}\right)\frac{2^{\ell+1-j}}{H}, 1\right\} \leqslant \min\left\{\frac{2^{\ell+4}}{H} + \frac{2^{2\ell+2-j}}{H^2}, 1\right\},$$

and this more than establishes (5.41) for every $\ell \geqslant 1$.

Step 5. For every $(m_2, m_3) \in \mathcal{Z}_j$, we use the identity

$$(m_2\cos\theta + m_3\sin\theta)^2 + (m_3\cos\theta - m_2\sin\theta)^2 = m_2^2 + m_3^2 \tag{5.44}$$

to study the product $\widetilde{\mathcal{I}}(m_2, m_3;\theta;u_1)\widetilde{\mathcal{I}}(m_3, -m_2;\theta;u_2)$.

Case 1: Assume that

$$\min\{(m_2\cos\theta + m_3\sin\theta)^2, (m_3\cos\theta - m_2\sin\theta)^2\} < 2^{j-3}. \tag{5.45}$$

Then

$$\begin{aligned}
&\max\{(m_2\cos\theta + m_3\sin\theta)^2, (m_3\cos\theta - m_2\sin\theta)^2\} \\
&\quad = (m_2^2 + m_3^2) - \min\{(m_2\cos\theta + m_3\sin\theta)^2, (m_3\cos\theta - m_2\sin\theta)^2\} \\
&\quad > 2^{2j-2} - 2^{j-3} > 2^{2j-3}.
\end{aligned} \tag{5.46}$$

Since $\mathbf{m}^* = (m_2, m_3) \in \mathcal{Z}_j$, it follows from (5.38) that the vector (m_2, m_3) has length greater than 2^{j-1}, while (5.45) shows that one of the two dot products

$$(m_2, m_3)\cdot(\cos\theta, \sin\theta) \quad \text{and} \quad (m_2, m_3)\cdot(-\sin\theta, \cos\theta)$$

has size less than $2^{(j-3)/2}$ which is substantially smaller than 2^{j-1} if j is large, meaning that the vector (m_2, m_3) is nearly perpendicular to one of the two vectors $(\cos\theta, \sin\theta)$ and $(-\sin\theta, \cos\theta)$ and nearly parallel to the other one, where nearly parallel means that the angle in between is in the range $2^{-j/2}$, so

$$\mathcal{Z}_j^\dagger = \{(m_2, m_3) \in \mathcal{Z}_j : (5.45) \text{ holds}\}$$

is the set of points in $\mathbb{Z}^2$ which fall into the dark shaded part in Figure 5.1, and we have generous upper bound

$$|\mathcal{Z}_j^\dagger| \leqslant 2^{10} 2^{2j} 2^{-j/2} = 2^{10+3j/2}. \tag{5.47}$$

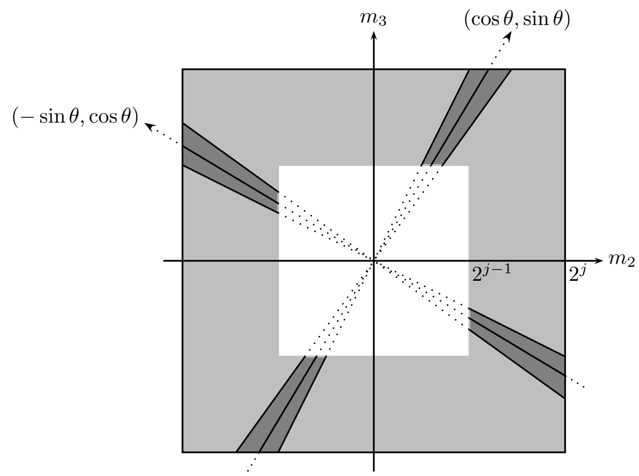

Figure 5.1: the location of $\mathbf{m}^* = (m_2, m_3) \in \mathcal{Z}_j^\dagger$

For every $\mathbf{m}^* = (m_2, m_3) \in \mathcal{Z}_j^\dagger$, it then follows from the bounds (5.33), (5.34) and (5.46) that

$$\begin{aligned}
&\widetilde{\mathcal{I}}(m_2, m_3; \theta; u_1)\widetilde{\mathcal{I}}(m_3, -m_2; \theta; u_2) \\
&\quad\leqslant \min\left\{u_1u_2, \frac{u_1}{\pi^2 u_2(m_3\cos\theta - m_2\sin\theta)^2}, \frac{u_2}{\pi^2 u_1(m_2\cos\theta + m_3\sin\theta)^2}\right\} \\
&\quad\leqslant \min\left\{u_1u_2, \max\left\{\frac{u_1}{u_2}, \frac{u_2}{u_1}\right\}\frac{1}{\pi^2 2^{2j-3}}\right\} \\
&\quad\leqslant \min\left\{u_1u_2, \frac{1}{2^{2j}}\max\left\{\frac{u_1}{u_2}, \frac{u_2}{u_1}\right\}\right\}.
\end{aligned} \tag{5.48}$$

Case 2: Suppose that (5.45) fails, so that $\mathbf{m}^* = (m_2, m_3) \in \mathcal{Z}_j^\ddagger = \mathcal{Z}_j \setminus \mathcal{Z}_j^\dagger$. Then

$$\min\{(m_2\cos\theta + m_3\sin\theta)^2, (m_3\cos\theta - m_2\sin\theta)^2\} \geqslant 2^{j-3}. \tag{5.49}$$

On the other hand, using the identity (5.44) and the inequality (5.38), we see that

$$\max\{(m_2\cos\theta + m_3\sin\theta)^2, (m_3\cos\theta - m_2\sin\theta)^2\} \geqslant \frac{m_2^2 + m_3^2}{2} \geqslant 2^{2j-3}. \tag{5.50}$$

It then follows from the bounds (5.33), (5.34), (5.49) and (5.50) that

$$\begin{aligned}
&\widetilde{\mathcal{I}}(m_2, m_3; \theta; u_1)\widetilde{\mathcal{I}}(m_3, -m_2; \theta; u_2) \\
&\quad\leqslant \min\left\{u_1u_2, \frac{1}{\pi^4 u_1u_2(m_2\cos\theta + m_3\sin\theta)^2(m_3\cos\theta - m_2\sin\theta)^2}\right\} \\
&\quad\leqslant \min\left\{u_1u_2, \frac{1}{\pi^4 u_1u_2 2^{3j-6}}\right\} \leqslant \min\left\{u_1u_2, \frac{1}{2^{3j}u_1u_2}\right\}.
\end{aligned} \tag{5.51}$$

Step 6. Using (5.32) and taking into account that the identity (5.36) holds whenever $\mathbf{m}^* = (m_2, m_3) = (0, 0)$, we now have the trivial upper bound

$$\int_0^1 \int_0^1 \Xi(\alpha_1, \alpha_2)\, \mathrm{d}\alpha_1\, \mathrm{d}\alpha_2 \leqslant \mathfrak{I}_1 + \mathfrak{I}_2, \tag{5.52}$$

where

$$\mathfrak{I}_1 = \sum_{j=0}^{\infty} \sum_{\ell=0}^{\infty} \sum_{\mathbf{m}^* \in \mathcal{Z}_j^\dagger} \int_{\Omega_j(\mathbf{m}^*;\ell)} \Lambda(\theta; u_1, u_2; H; \mathbf{v}; \mathbf{m})\, \mathrm{d}\alpha_1\, \mathrm{d}\alpha_2, \tag{5.53}$$

$$\mathfrak{I}_2 = \sum_{j=0}^{\infty} \sum_{\ell=0}^{\infty} \sum_{\mathbf{m}^* \in \mathcal{Z}_j^\ddagger} \int_{\Omega_j(\mathbf{m}^*;\ell)} \Lambda(\theta; u_1, u_2; H; \mathbf{v}; \mathbf{m})\, \mathrm{d}\alpha_1\, \mathrm{d}\alpha_2. \tag{5.54}$$

Lemma 5.2. *We have*

$$\mathfrak{I}_1 \leqslant 2^{23} \max\left\{\frac{u_1}{u_2^{1/2}}, \frac{u_2}{u_1^{1/2}}\right\}. \tag{5.55}$$

Proof. (i) Suppose first that $\ell = 0$. It follows from (5.31), (5.35) and (5.48) that for every $\mathbf{m}^* = (m_2, m_3) \in \mathcal{Z}_j^\dagger$ and $(\alpha_1, \alpha_2) \in \Omega_j(\mathbf{m}^*; 0)$, we have

$$0 \leqslant \Lambda(\theta; u_1, u_2; H; \mathbf{v}; \mathbf{m}) \leqslant \frac{H}{2} \min\left\{u_1 u_2, \frac{1}{2^{2j}} \max\left\{\frac{u_1}{u_2}, \frac{u_2}{u_1}\right\}\right\}.$$

Combining this with (5.41) and (5.47), we deduce that

$$\begin{aligned}
&\sum_{\mathbf{m}^* \in \mathcal{Z}_j^\dagger} \int_{\Omega_j(\mathbf{m}^*;0)} \Lambda(\theta; u_1, u_2; H; \mathbf{v}; \mathbf{m})\, \mathrm{d}\alpha_1\, \mathrm{d}\alpha_2 \\
&\qquad \leqslant 2^{10+3j/2} \lambda_2(\Omega_j(\mathbf{m}^*; 0)) \frac{H}{2} \min\left\{u_1 u_2, \frac{1}{2^{2j}} \max\left\{\frac{u_1}{u_2}, \frac{u_2}{u_1}\right\}\right\} \\
&\qquad \leqslant 2^{15+3j/2} \min\left\{u_1 u_2, \frac{1}{2^{2j}} \max\left\{\frac{u_1}{u_2}, \frac{u_2}{u_1}\right\}\right\}.
\end{aligned} \tag{5.56}$$

We claim that

$$\sum_{j=0}^{\infty} 2^{3j/2} \min\left\{u_1 u_2, \frac{1}{2^{2j}} \max\left\{\frac{u_1}{u_2}, \frac{u_2}{u_1}\right\}\right\} \leqslant 2^3 \max\left\{\frac{u_1}{u_2^{1/2}}, \frac{u_2}{u_1^{1/2}}\right\}. \tag{5.57}$$

To see this, let $J^\dagger$ denote the largest non-negative integer j such that

$$u_1 u_2 \leqslant \frac{1}{2^{2j}} \max\left\{\frac{u_1}{u_2}, \frac{u_2}{u_1}\right\}, \tag{5.58}$$

so that

$$2^{2J^\dagger} \leqslant \frac{1}{u_1 u_2} \max\left\{\frac{u_1}{u_2}, \frac{u_2}{u_1}\right\} \quad \text{and} \quad 2^{2J^\dagger+2} > \frac{1}{u_1 u_2} \max\left\{\frac{u_1}{u_2}, \frac{u_2}{u_1}\right\}. \tag{5.59}$$

Then it follows from (5.58) and (5.59) that

$$\begin{aligned}
&\sum_{j=0}^{J^\dagger} 2^{3j/2} \min\left\{u_1 u_2, \frac{1}{2^{2j}} \max\left\{\frac{u_1}{u_2}, \frac{u_2}{u_1}\right\}\right\} \leqslant u_1 u_2 \sum_{j=0}^{J^\dagger} 2^{3j/2} \\
&\qquad \leqslant 2 u_1 u_2 \left(\frac{1}{u_1 u_2} \max\left\{\frac{u_1}{u_2}, \frac{u_2}{u_1}\right\}\right)^{3/4} = 2 \max\left\{\frac{u_1}{u_2^{1/2}}, \frac{u_2}{u_1^{1/2}}\right\}
\end{aligned} \tag{5.60}$$

and

$$\sum_{j=J^\dagger+1}^{\infty} 2^{3j/2}\min\left\{u_1u_2,\frac{1}{2^{2j}}\max\left\{\frac{u_1}{u_2},\frac{u_2}{u_1}\right\}\right\}\leqslant\max\left\{\frac{u_1}{u_2},\frac{u_2}{u_1}\right\}\sum_{j=J^\dagger+1}^{\infty}2^{-j/2}$$

$$\leqslant 4\max\left\{\frac{u_1}{u_2},\frac{u_2}{u_1}\right\}\left(\frac{1}{u_1u_2}\max\left\{\frac{u_1}{u_2},\frac{u_2}{u_1}\right\}\right)^{-1/4}=4\max\left\{\frac{u_1}{u_2^{1/2}},\frac{u_2}{u_1^{1/2}}\right\}. \quad (5.61)$$

The inequality (5.57) follows on combining (5.60) and (5.61). Combining (5.56) and (5.57) now leads to the inequality

$$\sum_{j=0}^{\infty}\sum_{\mathbf{m}^*\in\mathcal{Z}_j^\dagger}\int_{\Omega_j(\mathbf{m}^*;0)}\Lambda(\theta;u_1,u_2;H;\mathbf{v};\mathbf{m})\,\mathrm{d}\alpha_1\,\mathrm{d}\alpha_2\leqslant 2^{18}\max\left\{\frac{u_1}{u_2^{1/2}},\frac{u_2}{u_1^{1/2}}\right\}. \quad (5.62)$$

(ii) Suppose next that $\ell\geqslant 1$. It follows from (5.31), (5.35), (5.40) and (5.48) that for every $\mathbf{m}^*=(m_2,m_3)\in\mathcal{Z}_j^\dagger$ and $(\alpha_1,\alpha_2)\in\Omega_j(\mathbf{m}^*;\ell)$, we have

$$0\leqslant\Lambda(\theta;u_1,u_2;H;\mathbf{v};\mathbf{m})\leqslant\frac{H}{2^{2\ell}}\min\left\{u_1u_2,\frac{1}{2^{2j}}\max\left\{\frac{u_1}{u_2},\frac{u_2}{u_1}\right\}\right\}. \quad (5.63)$$

Combining (5.41), (5.47) and (5.63), we deduce that

$$\begin{aligned}&\sum_{\mathbf{m}^*\in\mathcal{Z}_j^\dagger}\int_{\Omega_j(\mathbf{m}^*;\ell)}\Lambda(\theta;u_1,u_2;H;\mathbf{v};\mathbf{m})\,\mathrm{d}\alpha_1\,\mathrm{d}\alpha_2\\ &\quad\leqslant 2^{10+3j/2}\lambda_2(\Omega_j(\mathbf{m}^*;\ell))\frac{H}{2^{2\ell}}\min\left\{u_1u_2,\frac{1}{2^{2j}}\max\left\{\frac{u_1}{u_2},\frac{u_2}{u_1}\right\}\right\}\\ &\quad\leqslant 2^{10+3j/2}\min\left\{\frac{1}{2^{\ell-8}},\frac{H}{2^{2\ell}}\right\}\min\left\{u_1u_2,\frac{1}{2^{2j}}\max\left\{\frac{u_1}{u_2},\frac{u_2}{u_1}\right\}\right\},\end{aligned}$$

so that

$$\begin{aligned}&\sum_{\ell=1}^{\infty}\sum_{\mathbf{m}^*\in\mathcal{Z}_j^\dagger}\int_{\Omega_j(\mathbf{m}^*;\ell)}\Lambda(\theta;u_1,u_2;H;\mathbf{v};\mathbf{m})\,\mathrm{d}\alpha_1\,\mathrm{d}\alpha_2\\ &\quad\leqslant 2^{10+3j/2}\min\left\{u_1u_2,\frac{1}{2^{2j}}\max\left\{\frac{u_1}{u_2},\frac{u_2}{u_1}\right\}\right\}\sum_{\ell=1}^{\infty}\min\left\{\frac{1}{2^{\ell-8}},\frac{H}{2^{2\ell}}\right\}\\ &\quad\leqslant 2^{19+3j/2}\min\left\{u_1u_2,\frac{1}{2^{2j}}\max\left\{\frac{u_1}{u_2},\frac{u_2}{u_1}\right\}\right\}. \quad (5.64)\end{aligned}$$

Combining (5.57) and (5.64) now leads to the inequality

$$\sum_{j=0}^{\infty}\sum_{\ell=1}^{\infty}\sum_{\mathbf{m}^*\in\mathcal{Z}_j^\dagger}\int_{\Omega_j(\mathbf{m}^*;\ell)}\Lambda(\theta;u_1,u_2;H;\mathbf{v};\mathbf{m})\,\mathrm{d}\alpha_1\,\mathrm{d}\alpha_2\leqslant 2^{22}\max\left\{\frac{u_1}{u_2^{1/2}},\frac{u_2}{u_1^{1/2}}\right\}. \quad (5.65)$$

The inequality (5.55) now follows on combining (5.53), (5.62) and (5.65). ☐

Lemma 5.3. *We have*

$$\mathfrak{I}_2\leqslant\frac{2^{16}}{u_1^{1/3}u_2^{1/3}}. \quad (5.66)$$

Proof. (i) Suppose first that $\ell=0$. It follows from (5.31), (5.35) and (5.51) that for every $\mathbf{m}^*=(m_2,m_3)\in\mathcal{Z}_j^\ddagger$ and $(\alpha_1,\alpha_2)\in\Omega_j(\mathbf{m}^*;0)$, we have

$$0\leqslant\Lambda(\theta;u_1,u_2;H;\mathbf{v};\mathbf{m})\leqslant\frac{H}{2}\min\left\{u_1u_2,\frac{1}{2^{3j}u_1u_2}\right\}. \quad (5.67)$$

Note that it follows from (5.38) that $|\mathcal{Z}_0^\ddagger| \leqslant 8$ and $|\mathcal{Z}_j^\ddagger| \leqslant 4^{j+1}$ for $j \geqslant 1$, and so

$$|\mathcal{Z}_j^\ddagger| \leqslant 4^{j+2} \tag{5.68}$$

for all $j \geqslant 0$. Combining this with (5.41) and (5.67), we deduce that

$$\begin{aligned}\sum_{\mathbf{m}^*\in\mathcal{Z}_j^\ddagger} \int_{\Omega_j(\mathbf{m}^*;0)} &\Lambda(\theta;u_1,u_2;H;\mathbf{v};\mathbf{m})\,\mathrm{d}\alpha_1\,\mathrm{d}\alpha_2 \\ &\leqslant 4^{j+2}\lambda_2(\Omega_j(\mathbf{m}^*;0))\frac{H}{2}\min\left\{u_1u_2,\frac{1}{2^{3j}u_1u_2}\right\} \\ &\leqslant 2^{2j+9}\min\left\{u_1u_2,\frac{1}{2^{3j}u_1u_2}\right\}.\end{aligned} \tag{5.69}$$

We claim that

$$\sum_{j=0}^{\infty} 2^{2j}\min\left\{u_1u_2,\frac{1}{2^{3j}u_1u_2}\right\} \leqslant \frac{4}{u_1^{1/3}u_2^{1/3}}. \tag{5.70}$$

To see this, let $J^\ddagger$ denote the largest non-negative integer j such that

$$u_1u_2 \leqslant \frac{1}{2^{3j}u_1u_2}, \tag{5.71}$$

so that

$$2^{3J^\ddagger} \leqslant \frac{1}{u_1^2u_2^2} \quad \text{and} \quad 2^{3J^\ddagger+3} > \frac{1}{u_1^2u_2^2}. \tag{5.72}$$

Then it follows from (5.71) and (5.72) that

$$\begin{aligned}\sum_{j=0}^{J^\ddagger} 2^{2j}\min\left\{u_1u_2,\frac{1}{2^{3j}u_1u_2}\right\} &\leqslant u_1u_2\sum_{j=0}^{J^\ddagger}2^{2j} \\ &\leqslant 2u_1u_2\left(\frac{1}{u_1^2u_2^2}\right)^{2/3} = \frac{2}{u_1^{1/3}u_2^{1/3}}\end{aligned} \tag{5.73}$$

and

$$\begin{aligned}\sum_{j=J^\ddagger+1}^{\infty} 2^{2j}\min\left\{u_1u_2,\frac{1}{2^{3j}u_1u_2}\right\} &\leqslant \frac{1}{u_1u_2}\sum_{j=J^\ddagger+1}^{\infty}\frac{1}{2^j} \\ &\leqslant \frac{2}{u_1u_2}(u_1^2u_2^2)^{1/3} = \frac{2}{u_1^{1/3}u_2^{1/3}}.\end{aligned} \tag{5.74}$$

Combining (5.73) and (5.74) leads to (5.70) which, together with (5.69), leads to the inequality

$$\sum_{j=0}^{\infty}\sum_{\mathbf{m}^*\in\mathcal{Z}_j^\ddagger}\int_{\Omega_j(\mathbf{m}^*;0)}\Lambda(\theta;u_1,u_2;H;\mathbf{v};\mathbf{m})\,\mathrm{d}\alpha_1\,\mathrm{d}\alpha_2 \leqslant \frac{2^{11}}{u_1^{1/3}u_2^{1/3}}. \tag{5.75}$$

(ii) Suppose next that $\ell \geqslant 1$. It follows from (5.31), (5.35), (5.40) and (5.51) that for every $\mathbf{m}^* = (m_2,m_3) \in \mathcal{Z}_j^\ddagger$ and $(\alpha_1,\alpha_2) \in \Omega_j(\mathbf{m}^*;\ell)$, we have

$$0 \leqslant \Lambda(\theta;u_1,u_2;H;\mathbf{v};\mathbf{m}) \leqslant \frac{H}{2^{2\ell-2}}\min\left\{u_1u_2,\frac{1}{2^{3j}u_1u_2}\right\}. \tag{5.76}$$

Combining (5.41), (5.68) and (5.76), we deduce that

$$\begin{aligned}\sum_{\mathbf{m}^*\in\mathcal{Z}_j^\ddagger}&\int_{\Omega_j(\mathbf{m}^*;\ell)}\Lambda(\theta;u_1,u_2;H;\mathbf{v};\mathbf{m})\,\mathrm{d}\alpha_1\,\mathrm{d}\alpha_2\\&\leqslant 4^{j+2}\lambda_2(\Omega_j(\mathbf{m}^*;\ell))\frac{H}{2^{2\ell-2}}\min\left\{u_1u_2,\frac{1}{2^{3j}u_1u_2}\right\}\\&\leqslant 2^{2j+4}\min\left\{\frac{1}{2^{\ell-8}},\frac{H}{2^{2\ell-2}}\right\}\min\left\{u_1u_2,\frac{1}{2^{3j}u_1u_2}\right\},\end{aligned}$$

so that

$$\begin{aligned}\sum_{\ell=1}^{\infty}\sum_{\mathbf{m}^*\in\mathcal{Z}_j^\ddagger}&\int_{\Omega_j(\mathbf{m}^*;\ell)}\Lambda(\theta;u_1,u_2;H;\mathbf{v};\mathbf{m})\,\mathrm{d}\alpha_1\,\mathrm{d}\alpha_2\\&\leqslant 2^{2j+4}\min\left\{u_1u_2,\frac{1}{2^{3j}u_1u_2}\right\}\sum_{\ell=1}^{\infty}\min\left\{\frac{1}{2^{\ell-8}},\frac{H}{2^{2\ell-2}}\right\}\\&\leqslant 2^{2j+13}\min\left\{u_1u_2,\frac{1}{2^{3j}u_1u_2}\right\}.\end{aligned}\tag{5.77}$$

Combining (5.70) and (5.77) now leads to the inequality

$$\sum_{j=0}^{\infty}\sum_{\ell=1}^{\infty}\sum_{\mathbf{m}^*\in\mathcal{Z}_j^\ddagger}\int_{\Omega_j(\mathbf{m}^*;0)}\Lambda(\theta;u_1,u_2;H;\mathbf{v};\mathbf{m})\,\mathrm{d}\alpha_1\,\mathrm{d}\alpha_2\leqslant\frac{2^{15}}{u_1^{1/3}u_2^{1/3}}.\tag{5.78}$$

The inequality (5.66) now follows on combining (5.54), (5.75) and (5.78). ☐

It now follows from (5.52), (5.55) and (5.66) that

$$\int_0^1\int_0^1\Xi(\alpha_1,\alpha_2)\,\mathrm{d}\alpha_1\,\mathrm{d}\alpha_2\leqslant\psi(u_1,u_2),$$

where $\psi(u_1,u_2)$ is given by (5.1). Then for every real parameter $\kappa>1$, we have

$$\lambda_2(\{(\alpha_1,\alpha_2)\in[0,1]^2:\Xi(\alpha_1,\alpha_2)\geqslant\kappa\psi(u_1,u_2)\})\leqslant\frac{1}{\kappa}.\tag{5.79}$$

Step 7. It remains to justify all the steps. Choosing any real parameter $\kappa>1$, we remove the set

$$\mathfrak{B}(\kappa)=\{(\alpha_1,\alpha_2)\in[0,1]^2:\Xi(\alpha_1,\alpha_2)\geqslant\kappa\psi(u_1,u_2)\}$$

of κ-bad direction vectors $\mathbf{v}_0=(\alpha_1,\alpha_2)\in[0,1]^2$. In view of the estimate (5.79), the remaining set

$$\mathfrak{G}(\kappa)=[0,1]^2\setminus\mathfrak{B}(\kappa)$$

of κ-good direction vectors satisfies

$$\lambda_2(\mathfrak{G}(\kappa))\geqslant 1-\frac{1}{\kappa}.$$

Furthermore, for every $\mathbf{v}_0=(\alpha_1,\alpha_2)\in\mathfrak{G}(\kappa)$, we have the identity

$$\begin{aligned}\frac{2}{H}&\int_0^{H/2}\int_0^1\int_0^1\left(\sum_{\mathbf{n}\in\mathbb{Z}^3}\chi_{B(\gamma_1,\gamma_2,h)}(M\mathbf{n})\right)\mathrm{d}\gamma_1\,\mathrm{d}\gamma_2\,\mathrm{d}h\\&=\frac{2}{H}\int_0^{H/2}\int_0^1\int_0^1\left(\sum_{\mathbf{m}\in\mathbb{Z}^3}\int_{B(\gamma_1,\gamma_2,h)}\mathrm{e}^{-2\pi\mathrm{i}M^{-1}\mathbf{z}\cdot\mathbf{m}}\,\mathrm{d}\mathbf{z}\right)\mathrm{d}\gamma_1\,\mathrm{d}\gamma_2\,\mathrm{d}h,\end{aligned}\tag{5.80}$$

and the quantitative argument in Steps 4–6 shows that the infinite sum on the right hand side of (5.80) is absolutely convergent. This justifies the use of the Poisson summation formula and makes the argument precise.

This completes the proof of Lemma 5.1. □

6. Piecewise smooth barriers

Recall that in Theorems 1 and 4, we have made a rather specific restriction on the 2-coloring on the yz-parallel square faces of the n-cube 3-manifold, that each of the red and green parts is the union of finitely many polygons. Here we investigate how any polygon of the 2-colorings can be replaced by a circle, an ellipse, or any other piecewise smooth closed curve.

Theorem 5. *Let $n \geqslant 2$ be an integer, and let $\mathcal{M}$ be any n-cube 3-manifold with barriers, where each yz-parallel square face has a 2-coloring such that each of the red and green parts is the union of finitely many regions with edges that are piecewise smooth closed curves, satisfying some mild technical requirements as stated in Lemma 6.1, and where the green part has positive area. Suppose further that the Restriction on Red Coloring holds, and that there is a local repetition color-split neighborhood on the yz-parallel square faces. Then for almost every starting point and almost every direction $\mathbf{v} = (1, \alpha_1, \alpha_2) \in \mathbb{R}^3$, the corresponding half-infinite 1-direction geodesic is equidistributed in $\mathcal{M}$.*

Since the edges of a polygon have zero curvature, the buffer zone $\mathcal{B}_N$ within the local repetition color-split neighborhood which is central to our argument can be taken as a long and narrow rectangular color-split strip. To establish Theorem 5, it is sufficient to deal with some buffer zone color-split strip which involves *proper bending*, with non-zero curvature. We therefore need to establish a suitable analog of Lemma 3.4. Clearly it suffices to establish a suitable analog of Lemma 5.1.

The first step is to describe such a strip with proper bending.

Let $f : [a, b] \to \mathbb{R}$ be a monotonic and sufficiently smooth function. More precisely, we require that its derivative f' satisfies

$$\min_{x \in [a,b]} |f'(x)| > 0,$$

the curvature is non-zero in $[a, b]$, together with some mild technical requirements on higher order derivatives which we shall describe later.

The strip with proper bending is then a u-neighborhood

$$\mathcal{C}(f; u) = \mathcal{C}(f; a, b; u) = \{(x, y) \in \mathbb{R}^2 : x \in [a, b] \text{ and } y \in [f(x) - u, f(x) + u]\} \quad (6.1)$$

of a color-split curved edge

$$\mathcal{C}(f) = \mathcal{C}(f; a, b) = \{(x, y) \in \mathbb{R}^2 : x \in [a, b] \text{ and } y = f(x)\},$$

as illustrated in Figure 6.1.

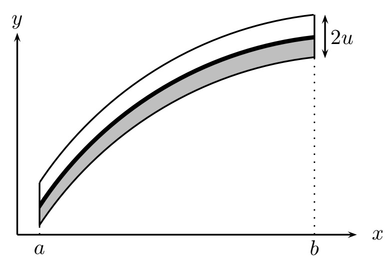

Figure 6.1: the u-neighborhood $\mathcal{C}(f; u)$ of the color-split curved edge $\mathcal{C}(f)$

Clearly the u-neighborhood $\mathcal{C}(f;u)$ has area

$$\lambda_2(\mathcal{C}(f;u)) = 2u(b-a).$$

We consider functions $f:[a,b]\to\mathbb{R}$ that are 3-times continuously differentiable, with positive constants $c_6,\ldots,c_{11}$, depending at most on the function f in the interval $[a,b]$, such that

$$0 < c_6 = \min_{x\in[a,b]} |f'(x)| \leqslant \max_{x\in[a,b]} |f'(x)| = c_7, \tag{6.2}$$

$$0 < c_8 = \min_{x\in[a,b]} |f''(x)| \leqslant \max_{x\in[a,b]} |f''(x)| = c_9, \tag{6.3}$$

$$0 < c_{10} = \min_{x\in[a,b]} |f'''(x)| \leqslant \max_{x\in[a,b]} |f'''(x)| = c_{11}. \tag{6.4}$$

We also require the function f to satisfy a mild technical condition, that there exists a constant $c_{12}=c_{12}(f;a,b)>0$, depending at most on the function f in the interval $[a,b]$, such that for any integer pair $(m_2,m_3)\in\mathbb{Z}^2\setminus\{(0,0)\}$, the interval $[a,b]$ is the union of at most c_{12} subintervals such that the functions

$$\frac{(m_2+m_3f'(x))^2}{f''(x)} \quad\text{and}\quad \frac{(m_2+m_3f'(x))^3}{f'(x)f''(x)} \tag{6.5}$$

are monotonic in each of the subintervals.

Lemma 6.1. *Let $f:[a,b]\to\mathbb{R}$ be a 3-times continuously differentiable function for which the conditions* (6.2)–(6.4) *as well as the technical condition concerning the functions* (6.5) *hold. For any even positive integer H, let*

$$G(\mathbf{s}_0;\alpha_1,\alpha_2;H) = |\{j=0,1,\ldots,H-1 : \mathbf{s}_0 + j(\alpha_1,\alpha_2)\in\mathcal{C}(f;u)\}|,$$

where the u-neighborhood $\mathcal{C}(f;u)$ is defined by (6.1). *Then for any parameter $\kappa>1$, we have*

$$\lambda_2\left(\left\{(\alpha_1,\alpha_2)\in[0,1]^2 : G(\mathbf{s}_0;\alpha_1,\alpha_2;H) \geqslant \frac{u(b-a)H}{4} - \kappa\psi(u)\right\}\right) \geqslant 1-\frac{1}{\kappa},$$

where

$$\psi(u) = c_{13}\left(u^{-1/2}+c_{14}\right), \tag{6.6}$$

and the constants $c_{13}=c_{13}(f;a,b)>0$ and $c_{14}=c_{14}(f;a,b)>0$ depend at most on the function f in the interval $[a,b]$.

Remark. Note that $(b-a)/2$ and u play the roles of u_1 and u_2 in Lemma 5.1, and that $G(\mathbf{s}_0;\alpha_1,\alpha_2;H)$ is the analog of the counting function $F(\mathbf{s}_0;\alpha_1,\alpha_2;H)$ there. We also have non-zero curvature instead of rotation there.

Proof of Lemma 6.1. We proceed by a number of steps corresponding to those in the proof of Lemma 5.1. We also use similar notation as much as possible.

Step 1. Here we give a good description of $G(\mathbf{s}_0;\alpha_1,\alpha_2;H)$. As before, we consider the lattice $\mathcal{L}(\alpha_1,\alpha_2)$ defined by (5.3) and (5.4). Then writing

$$B = (\mathcal{C}(f;u)-\mathbf{s}_0)\times[0,H)\subset\mathbb{R}^3,$$

we then have

$$G(\mathbf{s}_0;\alpha_1,\alpha_2;H) = |\mathcal{L}(\alpha_1,\alpha_2)\cap B| = \sum_{\mathbf{n}\in\mathbb{Z}^3}\chi_B(M\mathbf{n}).$$

Applying the Poisson summation formula (5.6) with $f=\chi_B$, we have, analogous to (5.7), the formal identity

$$G(\mathbf{s}_0;\alpha_1,\alpha_2;H) = \sum_{\mathbf{m}\in\mathbb{Z}^3}\int_B \mathrm{e}^{-2\pi\mathrm{i}M^{-1}\mathbf{z}\cdot\mathbf{m}}\,\mathrm{d}\mathbf{z}. \tag{6.7}$$

Note, in particular, that for $\mathbf{m} = \mathbf{0}$, we have

$$\int_B \mathrm{e}^{-2\pi\mathrm{i}M^{-1}\mathbf{z}\cdot\mathbf{m}}\,\mathrm{d}\mathbf{z} = \lambda_3(B) = \lambda_2(\mathcal{C}(f;u))H = 2u(b-a)H. \tag{6.8}$$

Then, using (5.9) and analogous to (5.10)–(5.12), for every $\mathbf{m} \in \mathbb{Z}^3 \setminus \{\mathbf{0}\}$, we can write

$$\int_B \mathrm{e}^{-2\pi\mathrm{i}M^{-1}\mathbf{z}\cdot\mathbf{m}}\,\mathrm{d}\mathbf{z} = \mathcal{I}(\mathbf{m};\mathcal{C}(f;u)-\mathbf{s}_0)\mathcal{J}(\mathbf{m};\mathbf{v};H), \tag{6.9}$$

where

$$\mathcal{I}(\mathbf{m};\mathcal{C}(f;u)-\mathbf{s}_0) = \int_{\mathcal{C}(f;u)-\mathbf{s}_0} \mathrm{e}^{2\pi\mathrm{i}(z_1m_2+z_2m_3)}\,\mathrm{d}z_1\,\mathrm{d}z_2 \tag{6.10}$$

and

$$\mathcal{J}(\mathbf{m};\mathbf{v};H) = \int_0^H \mathrm{e}^{-2\pi\mathrm{i}z_3\mathbf{v}\cdot\mathbf{m}}\,\mathrm{d}z_3 = \frac{1-\mathrm{e}^{-2\pi\mathrm{i}H\mathbf{v}\cdot\mathbf{m}}}{2\pi\mathrm{i}\mathbf{v}\cdot\mathbf{m}}. \tag{6.11}$$

To study the term (6.10), note from (6.1) that a typical point in $\mathcal{C}(f;u)$ is of the form $(x, f(x)+t)$, where $x \in [a,b]$ and $-u \leqslant t \leqslant u$. We therefore use the substitution

$$(z_1, z_2) = (x, f(x)+t) - \mathbf{s}_0, \quad \text{with Jacobian} \quad f'(x),$$

so that, writing $\mathbf{m}^* = (m_2, m_3)$, we have

$$\begin{aligned}
\mathcal{I}(\mathbf{m};\mathcal{C}(f;u)-\mathbf{s}_0) &= \int_a^b\int_{-u}^u \mathrm{e}^{2\pi\mathrm{i}((x,f(x)+t)-\mathbf{s}_0)\cdot\mathbf{m}^*} f'(x)\,\mathrm{d}t\,\mathrm{d}x \\
&= \mathrm{e}^{-2\pi\mathrm{i}\mathbf{s}_0\cdot\mathbf{m}^*}\int_a^b\int_{-u}^u \mathrm{e}^{2\pi\mathrm{i}(x,f(x)+t)\cdot\mathbf{m}^*} f'(x)\,\mathrm{d}t\,\mathrm{d}x \\
&= \mathrm{e}^{-2\pi\mathrm{i}\mathbf{s}_0\cdot\mathbf{m}^*}\mathcal{I}_1(\mathbf{m};\mathcal{C}(f;u))\mathcal{I}_2(\mathbf{m};\mathcal{C}(f;u)),
\end{aligned} \tag{6.12}$$

where

$$\mathcal{I}_1(\mathbf{m};\mathcal{C}(f;u)) = \int_a^b \mathrm{e}^{2\pi\mathrm{i}(m_2x+m_3f(x))} f'(x)\,\mathrm{d}x \tag{6.13}$$

and

$$\mathcal{I}_2(\mathbf{m};\mathcal{C}(f;u)) = \int_{-u}^u \mathrm{e}^{2\pi\mathrm{i}tm_3}\,\mathrm{d}t = \frac{\sin(2\pi u m_3)}{\pi m_3}, \tag{6.14}$$

provided that $m_3 \neq 0$. If $m_3 = 0$, then clearly

$$\mathcal{I}_2(\mathbf{m};\mathcal{C}(f;u)) = \int_{-u}^u \mathrm{d}t = 2u. \tag{6.15}$$

The analysis of the integral (6.13) is complicated. However, the smoothness of the function f makes it possible to have effective estimates. We observe that it is possible to establish good estimates for complex exponential integrals whenever the exponential function exhibits *rapid fluctuations*. The worst case scenario is when the exponential function $\mathrm{e}^{2\pi\mathrm{i}(m_2x+m_3f(x))}$ is almost constant in a small neighborhood of some point $x_0 \in [a,b]$.

Let us consider some heuristics. Write $h(x) = m_2x + m_3f(x)$. Then we have the finite Taylor expansion

$$\begin{aligned}
h(x) - h(x_0) &= m_2(x-x_0) + m_3(f(x)-f(x_0)) \\
&= (x-x_0)(m_2+m_3f'(x_0)) + \frac{(x-x_0)^2m_3f''(x_0)}{2} + \frac{(x-x_0)^3m_3f'''(y)}{6}
\end{aligned} \tag{6.16}$$

for some appropriate y between x_0 and x. Suppose further that there exists some $x_0 \in [a,b]$ such that $h'(x_0) = m_2 + m_3 f'(x_0) = 0$. Then for every $x \in [a,b]$, the Taylor expansion (6.16) simplifies to

$$h(x) - h(x_0) = \frac{(x-x_0)^2 m_3 f''(x_0)}{2} + \frac{(x-x_0)^3 m_3 f'''(y)}{6} \tag{6.17}$$

for some $y \in [a,b]$. The Taylor expansions (6.16) and (6.17) show that the complex exponential function

$$\mathrm{e}^{2\pi\mathrm{i}h(x)} = \mathrm{e}^{2\pi\mathrm{i}(m_2x+m_3f(x))}$$

starts to exhibit increasing fluctuations if x is *relatively far* from x_0.

The estimation of *fluctuating integrals* is a well known general problem in number theory and especially in the theory of the Riemann zeta-function. The following result can be found in the treatise of Titchmarsh [5, Lemma 4.3] on the latter.

Lemma 6.2. *Suppose that for real functions F and G, the quotient $F'(x)/G(x)$ is monotonic in the interval $[a,b]$, and $|F'(x)/G(x)| \geqslant \mu > 0$ for every $x \in [a,b]$. Then*

$$\left| \int_a^b \mathrm{e}^{\mathrm{i}F(x)} G(x)\,\mathrm{d}x \right| \leqslant \frac{4}{\mu}.$$

To estimate the integral (6.13), we consider the functions

$$F(x) = 2\pi(m_2x + m_3f(x)) \quad \text{and} \quad G(x) = f'(x)$$

in the interval $[a,b]$. Note that

$$\frac{F'(x)}{G(x)} = \frac{2\pi(m_2 + m_3f'(x))}{f'(x)} = \frac{2\pi m_2}{f'(x)} + 2\pi m_3, \tag{6.18}$$

and the derivative

$$\frac{\mathrm{d}}{\mathrm{d}x}\frac{F'(x)}{G(x)} = -\frac{2\pi m_2 f''(x)}{(f'(x))^2}$$

has constant sign in $[a,b]$, since $f'(x)$ and $f''(x)$ are continuous and non-zero in $[a,b]$, in view of (6.2) and (6.3). Hence the function (6.18) is monotonic in $[a,b]$. To apply Lemma 6.2, we need to ensure that the function (6.18) is also bounded away from 0. If no x_0 is in or near the interval $[a,b]$ such that $F'(x_0) = 2\pi(m_2 + m_3f'(x_0)) = 0$, then we can find a good value for μ in Lemma 6.2. However, we clearly have a problem if there exists $x_0 \in [a,b]$ such that $F'(x_0) = 2\pi(m_2 + m_3f'(x_0)) = 0$.

The following is the worst case scenario. Suppose that $x_0 \in (a,b)$. Let $\delta > 0$ be a *small* parameter, to be specified later, such that $a < x_0 - \delta < x_0 + \delta < b$. Then

$$\begin{aligned} \mathcal{I}_1(\mathbf{m};\mathcal{C}(f;u)) &= \int_a^b \mathrm{e}^{2\pi\mathrm{i}(m_2x+m_3f(x))} f'(x)\,\mathrm{d}x \\ &= \mathcal{I}_1^{(-)}(\mathbf{m};\mathcal{C}(f;u)) + \mathcal{I}_1^{(0)}(\mathbf{m};\mathcal{C}(f;u)) + \mathcal{I}_1^{(+)}(\mathbf{m};\mathcal{C}(f;u)), \end{aligned} \tag{6.19}$$

where

$$\mathcal{I}_1^{(-)}(\mathbf{m};\mathcal{C}(f;u)) = \int_a^{x_0-\delta} \mathrm{e}^{2\pi\mathrm{i}(m_2x+m_3f(x))} f'(x)\,\mathrm{d}x,$$

$$\mathcal{I}_1^{(0)}(\mathbf{m};\mathcal{C}(f;u)) = \int_{x_0-\delta}^{x_0+\delta} \mathrm{e}^{2\pi\mathrm{i}(m_2x+m_3f(x))} f'(x)\,\mathrm{d}x,$$

$$\mathcal{I}_1^{(+)}(\mathbf{m};\mathcal{C}(f;u)) = \int_{x_0+\delta}^{b} \mathrm{e}^{2\pi\mathrm{i}(m_2x+m_3f(x))} f'(x)\,\mathrm{d}x.$$

For the integral $\mathcal{I}_1^{(0)}(\mathbf{m};\mathcal{C}(f;u))$, we have the trivial estimate

$$|\mathcal{I}_1^{(0)}(\mathbf{m};\mathcal{C}(f;u))| \leqslant 2\delta \max_{x\in[a,b]} |f'(x)| = 2c_7\delta, \tag{6.20}$$

in view of (6.2). For the integral $\mathcal{I}_1^{(+)}(\mathbf{m};\mathcal{C}(f;u))$, we use the Taylor expansion

$$\begin{aligned} F'(x_0+\delta) &= 2\pi(m_2 + m_3 f'(x_0+\delta)) \\ &= 2\pi(m_2+m_3f'(x_0)) + 2\pi m_3 \left(\delta f''(x_0) + \frac{\delta^2 f'''(y_+)}{2}\right) \\ &= 2\pi m_3 \left(\delta f''(x_0) + \frac{\delta^2 f'''(y_+)}{2}\right) \end{aligned}$$

for some $y_+ \in (x_0, x_0+\delta)$. Since $F'(x_0)/G(x_0) = 0$ and $F'(x)/G(x)$ is monotonic in $[a,b]$, it follows that for every $x \in [x_0+\delta, b]$, we have

$$\begin{aligned} \left|\frac{F'(x)}{G(x)}\right| \geqslant \left|\frac{F'(x_0+\delta)}{G(x_0+\delta)}\right| &= \frac{2\pi|m_3|}{|f'(x_0+\delta)|}\left|\delta f''(x_0) + \frac{\delta^2 f'''(y_+)}{2}\right| \\ &\geqslant \frac{2\pi|m_3|}{|f'(x_0+\delta)|}\left(\delta|f''(x_0)| - \frac{\delta^2|f'''(y_+)|}{2}\right). \end{aligned} \tag{6.21}$$

For the integral $\mathcal{I}_1^{(-)}(\mathbf{m};\mathcal{C}(f;u))$, we use the Taylor expansion

$$\begin{aligned} F'(x_0-\delta) &= 2\pi(m_2 + m_3 f'(x_0-\delta)) \\ &= 2\pi(m_2+m_3f'(x_0)) + 2\pi m_3 \left(-\delta f''(x_0) + \frac{\delta^2 f'''(y_-)}{2}\right) \\ &= 2\pi m_3 \left(-\delta f''(x_0) + \frac{\delta^2 f'''(y_-)}{2}\right) \end{aligned}$$

for some $y_- \in (x_0-\delta, x_0)$. Since $F'(x_0)/G(x_0) = 0$ and $F'(x)/G(x)$ is monotonic in $[a,b]$, it follows that for every $x \in [a, x_0-\delta]$, we have

$$\begin{aligned} \left|\frac{F'(x)}{G(x)}\right| \geqslant \left|\frac{F'(x_0-\delta)}{G(x_0-\delta)}\right| &= \frac{2\pi|m_3|}{|f'(x_0-\delta)|}\left|\delta f''(x_0) - \frac{\delta^2 f'''(y_-)}{2}\right| \\ &\geqslant \frac{2\pi|m_3|}{|f'(x_0-\delta)|}\left(\delta|f''(x_0)| - \frac{\delta^2|f'''(y_-)|}{2}\right). \end{aligned} \tag{6.22}$$

Combining (6.2)–(6.4), (6.21) and (6.22), we conclude that for any $x \in [a, x_0-\delta]$ or any $x \in [x_0+\delta, b]$, we have the lower bound

$$\left|\frac{F'(x)}{G(x)}\right| \geqslant \frac{2\pi|m_3|}{c_7}\left(c_8\delta - \frac{c_{11}\delta^2}{2}\right) \geqslant c_{15}\delta|m_3|,$$

provided that $0 < \delta \leqslant c_{16}$, where $c_{15} = c_{15}(f;a,b) > 0$ and $c_{16} = c_{16}(f;a,b) > 0$ are constants depending at most on the function f in the interval $[a,b]$. Using this bound, Lemma 6.2 then gives the estimates

$$|\mathcal{I}_1^{(\pm)}(\mathbf{m};\mathcal{C}(f;u))| \leqslant \frac{4}{c_{15}\delta|m_3|}. \tag{6.23}$$

Combining (6.19), (6.20) and (6.23), we arrive at the estimate

$$|\mathcal{I}_1(\mathbf{m};\mathcal{C}(f;u))| \leqslant 2c_7\delta + \frac{8}{c_{15}\delta|m_3|} \leqslant \frac{c_{17}}{|m_3|^{1/2}} \tag{6.24}$$

if we choose δ in the range of $|m_3|^{-1/2}$, where the constant $c_{17} = c_{17}(f;a,b) > 0$ depends at most on the function f in the interval $[a,b]$.

The above analysis and the estimate (6.24) are only valid provided that $m_3 \neq 0$. If $m_3 = 0$, then (6.13) becomes

$$\mathcal{I}_1(\mathbf{m}; \mathcal{C}(f;u)) = \int_a^b \mathrm{e}^{2\pi\mathrm{i}m_2 x} f'(x)\,\mathrm{d}x,$$

and it follows from (6.2) that $|\mathcal{I}_1(\mathbf{m}; \mathcal{C}(f;u))| \leqslant (b-a)c_7$. For convenience, we can choose $c_{17} = c_{17}(f;a,b) > 0$ sufficiently large so that

$$c_{17} \geqslant (b-a)c_7.$$

Then

$$|\mathcal{I}_1(\mathbf{m}; \mathcal{C}(f;u))| \leqslant c_{17}. \tag{6.25}$$

We emphasize again that the above represents the worst case scenario. Of course, the estimates (6.24) and (6.25) remain valid if $m_2 + m_3 f'(x) \neq 0$ for any $x \in [a,b]$.

Combining (6.7)–(6.9) and (6.12), we conclude that

$$\begin{aligned} &G(\mathbf{s}_0; \alpha_1, \alpha_2; H) - 2u(b-a)H \\ &\quad = \sum_{\mathbf{m}\in\mathbb{Z}^3\setminus\{\mathbf{0}\}} \mathrm{e}^{-2\pi\mathrm{i}\mathbf{s}_0\cdot\mathbf{m}^*} \mathcal{I}_1(\mathbf{m}; \mathcal{C}(f;u)) \mathcal{I}_2(\mathbf{m}; \mathcal{C}(f;u)) \mathcal{J}(\mathbf{m}; \mathbf{v}; H), \end{aligned}$$

where estimates for the various factors in the summand are given by (6.11), (6.14), (6.15), (6.24) and (6.25).

Step 2. Here we mimic Step 2 in Section 5, contract the interval $[0,H]$ in the third direction and average over all contractions in a similar manner. We then conclude, analogous to (5.23), that

$$\begin{aligned} &G(\mathbf{s}_0; \alpha_1, \alpha_2; H) - u(b-a)H \\ &\quad \geqslant \sum_{\mathbf{m}\in\mathbb{Z}^3\setminus\{\mathbf{0}\}} \mathrm{e}^{-2\pi\mathrm{i}\mathbf{s}_0\cdot\mathbf{m}^*} \mathcal{I}_1(\mathbf{m}; \mathcal{C}(f;u)) \mathcal{I}_2(\mathbf{m}; \mathcal{C}(f;u)) \widetilde{\mathcal{J}}(\mathbf{m}; \mathbf{v}; H), \end{aligned}$$

where the factor $\widetilde{\mathcal{J}}(\mathbf{m}; \mathbf{v}; H)$ is given by (5.22).

Step 3. Here we also contract the intervals $[a,b]$ and $[f(x)-u, f(x)+u]$ in $\mathcal{C}(f;u)$ and average over all contractions. More precisely, for every γ_1 and γ_2 satisfying $0 \leqslant \gamma_1 \leqslant (b-a)/2$ and $0 \leqslant \gamma_2 \leqslant 1$, consider the smaller set

$$\begin{aligned} \Delta(\gamma_1,\gamma_2) &= \mathcal{C}(f; a+\gamma_1, b-\gamma_1; \gamma_2 u) \\ &= \{(x,y) \in \mathbb{R}^2 : x \in [a+\gamma_1, b-\gamma_1] \text{ and } y \in [f(x)-\gamma_2 u, f(x)+\gamma_2 u]\}. \end{aligned}$$

Then, analogous to (5.24) and (5.25), we have

$$\begin{aligned} &G(\mathbf{s}_0; \alpha_1, \alpha_2; H) - \frac{u(b-a)H}{4} \\ &\geqslant \frac{4}{(b-a)H} \int_0^{H/2} \int_0^1 \int_0^{(b-a)/2} \left(\sum_{\mathbf{m}\in\mathbb{Z}^3\setminus\{\mathbf{0}\}} \int_{B(\gamma_1,\gamma_2,h)} \mathrm{e}^{-2\pi\mathrm{i}M^{-1}\mathbf{z}\cdot\mathbf{m}}\,\mathrm{d}\mathbf{z} \right) \mathrm{d}\gamma_1\,\mathrm{d}\gamma_2\,\mathrm{d}h, \end{aligned} \tag{6.26}$$

where

$$B(\gamma_1,\gamma_2,h) = (\Delta(\gamma_1,\gamma_2) - \mathbf{s}_0) \times [h, H-h),$$

and, analogous to (5.26), we have

$$\int_{B(\gamma_1,\gamma_2,h)} \mathrm{e}^{-2\pi\mathrm{i}M^{-1}\mathbf{z}\cdot\mathbf{m}}\,\mathrm{d}\mathbf{z} = \mathcal{I}(\mathbf{m}; \Delta(\gamma_1,\gamma_2) - \mathbf{s}_0) \mathcal{J}(\mathbf{m}; \mathbf{v}; H; h). \tag{6.27}$$

Here $\mathcal{J}(\mathbf{m};\mathbf{v};H;h)$ is given by (5.21) and

$$\mathcal{I}(\mathbf{m};\Delta(\gamma_1,\gamma_2)-\mathbf{s}_0)=\mathrm{e}^{-2\pi\mathrm{i}\mathbf{s}_0\cdot\mathbf{m}^*}\mathcal{I}_1(\mathbf{m};\Delta(\gamma_1,\gamma_2))\mathcal{I}_2(\mathbf{m};\Delta(\gamma_1,\gamma_2)), \tag{6.28}$$

where

$$\mathcal{I}_1(\mathbf{m};\Delta(\gamma_1,\gamma_2))=\int_{a+\gamma_1}^{b-\gamma_1}\mathrm{e}^{2\pi\mathrm{i}(m_2x+m_3f(x))}f'(x)\,\mathrm{d}x \tag{6.29}$$

and

$$\mathcal{I}_2(\mathbf{m};\Delta(\gamma_1,\gamma_2))=\int_{-\gamma_2u}^{\gamma_2u}\mathrm{e}^{2\pi\mathrm{i}tm_3}\,\mathrm{d}t=\frac{\sin(2\pi\gamma_2um_3)}{\pi m_3}, \tag{6.30}$$

provided that $m_3\neq 0$. If $m_3=0$, then clearly

$$\mathcal{I}_2(\mathbf{m};\Delta(\gamma_1,\gamma_2))=2\gamma_2u. \tag{6.31}$$

Write

$$\widetilde{\mathcal{I}}_1(\mathbf{m};\mathcal{C}(f;u))=\frac{2}{b-a}\int_0^{(b-a)/2}\int_{a+\gamma_1}^{b-\gamma_1}\mathrm{e}^{2\pi\mathrm{i}(m_2x+m_3f(x))}f'(x)\,\mathrm{d}x\,\mathrm{d}\gamma_1 \tag{6.32}$$

and

$$\begin{aligned}\widetilde{\mathcal{I}}_2(\mathbf{m};\mathcal{C}(f;u))&=\int_0^1\int_{-\gamma_2u}^{\gamma_2u}\mathrm{e}^{2\pi\mathrm{i}tm_3}\,\mathrm{d}t\,\mathrm{d}\gamma_2\\&=\int_0^1\frac{\sin(2\pi\gamma_2um_3)}{\pi m_3}\,\mathrm{d}\gamma_2=\frac{2\sin^2(\pi um_3)}{u(\pi m_3)^2},\end{aligned} \tag{6.33}$$

provided that $m_3\neq 0$. If $m_3=0$, then clearly

$$\widetilde{\mathcal{I}}_2(\mathbf{m};\mathcal{C}(f;u))=\int_0^1\int_{-\gamma_2u}^{\gamma_2u}\mathrm{d}t\,\mathrm{d}\gamma_2=\int_0^1 2\gamma_2u\,\mathrm{d}\gamma_2=u. \tag{6.34}$$

It now follows from (5.22) and (6.26)–(6.34) that

$$\begin{aligned}&G(\mathbf{s}_0;\alpha_1,\alpha_2;H)-\frac{u(b-a)H}{4}\\&\quad\geqslant\sum_{\mathbf{m}\in\mathbb{Z}^3\setminus\{\mathbf{0}\}}\mathrm{e}^{-2\pi\mathrm{i}\mathbf{s}_0\cdot\mathbf{m}^*}\widetilde{\mathcal{I}}_1(\mathbf{m};\mathcal{C}(f;u))\widetilde{\mathcal{I}}_2(\mathbf{m};\mathcal{C}(f;u))\widetilde{\mathcal{J}}(\mathbf{m};\mathbf{v};H).\end{aligned} \tag{6.35}$$

Step 4. It follows from (6.35) that

$$G(\mathbf{s}_0;\alpha_1,\alpha_2;H)\geqslant\frac{u(b-a)H}{4}+\sum_{\mathbf{m}\in\mathbb{Z}^3\setminus\{\mathbf{0}\}}\Lambda(\mathcal{C}(f;u);H;\mathbf{v};\mathbf{m}),$$

where

$$\Lambda(\mathcal{C}(f;u);H;\mathbf{v};\mathbf{m})=|\widetilde{\mathcal{I}}_1(\mathbf{m};\mathcal{C}(f;u))|\,|\widetilde{\mathcal{I}}_2(\mathbf{m};\mathcal{C}(f;u))|\,|\widetilde{\mathcal{J}}(\mathbf{m};\mathbf{v};H)|. \tag{6.36}$$

As before in Section 5, the inequalities in Steps 1–3 are at this stage only formal inequalities, as we have not considered the question of convergence. Write

$$\Xi(\alpha_1,\alpha_2)=\sum_{\mathbb{Z}^3\setminus\{\mathbf{0}\}}\Lambda(\mathcal{C}(f;u);H;\mathbf{v};\mathbf{m}), \tag{6.37}$$

where $\mathbf{v}=(1,\alpha_1,\alpha_2)$. We shall use the first-moment method and analyze the average

$$\int_0^1\int_0^1\Xi(\alpha_1,\alpha_2)\,\mathrm{d}\alpha_1\,\mathrm{d}\alpha_2,$$

and remove those directions (α_1,α_2) for which $\Xi(\alpha_1,\alpha_2)$ is substantially larger than the average. In this step, we do some further preparation.

First of all, note that the estimates (6.24) and (6.25) are independent of a and b. Combining these with (6.29) and (6.32), we obtain the bound

$$|\widetilde{\mathcal{I}}_1(\mathbf{m};\mathcal{C}(f;u))| \leqslant c_{17}\min\left\{\frac{1}{|m_3|^{1/2}},1\right\}. \tag{6.38}$$

On the other hand, using (6.30), (6.31) and the inequality $|\sin y| \leqslant \min\{|y|,1\}$, which holds for every $y\in\mathbb{R}$, we obtain the bound

$$|\widetilde{\mathcal{I}}_2(\mathbf{m};\mathcal{C}(f;u))| \leqslant 2\min\left\{u,\frac{1}{u(\pi m_3)^2}\right\}. \tag{6.39}$$

Meanwhile, the estimate for the term $|\widetilde{\mathcal{J}}(\mathbf{m};\mathbf{v};H)|$ is given by (5.35). Furthermore, since H is even, in view of (5.36), we may assume that $\mathbf{m}^*=(m_2,m_3)\in\mathbb{Z}^2\setminus\{(0,0)\}$.

Finally, the definitions and estimates (5.37)–(5.41) concerning $\mathcal{Z}_j$ and $\Omega_j(\mathbf{m}^*;\ell)$, where $j=0,1,2,3,\ldots$ and $\ell=0,1,2,3,\ldots$, remain valid.

Step 5. Recall our comment in Step 1 that the complex exponential function

$$\mathrm{e}^{2\pi\mathrm{i}h(x)}=\mathrm{e}^{2\pi\mathrm{i}(m_2x+m_3f(x))}$$

starts to exhibit increasing fluctuations if x is *relatively far* from any root x_0 of the equation $m_2+m_3f'(x)=0$. If x_0 is far from the interval $[a,b]$, then the lack of fluctuation near x_0 is not a problem. However, it is a serious problem if $x_0\in[a,b]$.

Motivated by this observation, we split the argument into two cases, the first of which and some of its consequences are summarized by the following lemma.

Lemma 6.3. *Suppose that $(m_2,m_3)\in\mathbb{Z}^2\setminus\{(0,0\}$ and the condition*

$$\min_{a\leqslant x\leqslant b}|m_2+m_3f'(x)|<\frac{1}{4}\max\{|m_2|,|m_3|\}\min\left\{\min_{a\leqslant x\leqslant b}|f'(x)|,1\right\} \tag{6.40}$$

is satisfied. Then

(i) *$m_2m_3\neq 0$; and*

(ii) *the inequality*

$$\frac{1}{c_{18}}\leqslant\left|\frac{m_2}{m_3}\right|\leqslant c_{18} \tag{6.41}$$

holds, where the constant $c_{18}=c_{18}(f;a,b)>1$ depends at most on the function f in the interval $[a,b]$.

Proof. (i) If $m_2=0$, then the inequality (6.40) becomes

$$|m_3|\min_{a\leqslant x\leqslant b}|f'(x)|<\frac{1}{4}|m_3|\min\left\{\min_{a\leqslant x\leqslant b}|f'(x)|,1\right\}$$

which is clearly absurd. If $m_3=0$, then the inequality (6.40) becomes

$$|m_2|<\frac{1}{4}|m_2|\min\left\{\min_{a\leqslant x\leqslant b}|f'(x)|,1\right\}$$

which is also clearly absurd. Thus $m_2m_3\neq 0$ as claimed.

(ii) Dividing both sides by $m_3\neq 0$, the inequality (6.40) becomes

$$\min_{a\leqslant x\leqslant b}\left|\frac{m_2}{m_3}+f'(x)\right|<\frac{1}{4}\max\left\{\left|\frac{m_2}{m_3}\right|,1\right\}\min\left\{\min_{a\leqslant x\leqslant b}|f'(x)|,1\right\}. \tag{6.42}$$

Suppose first that

$$\frac{m_2}{m_3}=C_L,$$

where $|C_L|$ is *large*. Then using (6.2), the left and right sides of (6.42) are

$$\min_{a\leqslant x\leqslant b}\left|\frac{m_2}{m_3}+f'(x)\right| = \min_{a\leqslant x\leqslant b}|C_L+f'(x)| \geqslant |C_L| - \max_{a\leqslant x\leqslant b}|f'(x)| \geqslant |C_L| - c_7$$

and

$$\frac{1}{4}\max\left\{\left|\frac{m_2}{m_3}\right|,1\right\}\min\left\{\min_{a\leqslant x\leqslant b}|f'(x)|,1\right\} = \frac{1}{4}|C_L|\min\left\{\min_{a\leqslant x\leqslant b}|f'(x)|,1\right\} \leqslant \frac{1}{4}|C_L|$$

respectively. Clearly

$$|C_L| - c_7 > \frac{1}{4}|C_L|$$

if $|C_L|$ is sufficiently large in terms of the function f in the interval $[a,b]$, so that (6.42) fails. This establishes the upper bound in (6.41). Suppose next that

$$\frac{m_2}{m_3} = C_S,$$

where $|C_S|$ is *small*. Then the left and right sides of (6.42) are

$$\min_{a\leqslant x\leqslant b}\left|\frac{m_2}{m_3}+f'(x)\right| = \min_{a\leqslant x\leqslant b}|C_S+f'(x)| \geqslant \min_{a\leqslant x\leqslant b}|f'(x)| - |C_S|$$

and

$$\frac{1}{4}\max\left\{\left|\frac{m_2}{m_3}\right|,1\right\}\min\left\{\min_{a\leqslant x\leqslant b}|f'(x)|,1\right\} = \frac{1}{4}\min\left\{\min_{a\leqslant x\leqslant b}|f'(x)|,1\right\} \leqslant \frac{1}{4}\min_{a\leqslant x\leqslant b}|f'(x)|$$

respectively. Clearly

$$\min_{a\leqslant x\leqslant b}|f'(x)| - |C_S| > \frac{1}{4}\min_{a\leqslant x\leqslant b}|f'(x)|$$

if $|C_S|$ is sufficiently small in terms of the function f in the interval $[a,b]$, so that (6.42) fails. This establishes the lower bound in (6.41). $\square$

We also need to consider the case when the condition (6.40) is violated, so that the opposite condition

$$\min_{a\leqslant x\leqslant b}|m_2+m_3f'(x)| \geqslant \frac{1}{4}\max\{|m_2|,|m_3|\}\min\left\{\min_{a\leqslant x\leqslant b}|f'(x)|,1\right\} \tag{6.43}$$

is satisfied.

Accordingly, for every $j=1,2,3,\ldots$, write $\mathcal{Z}_j = \mathcal{Z}_j^\dagger \cup \mathcal{Z}_j^\ddagger$, where

$$\mathcal{Z}_j^\dagger = \{(m_2,m_3)\in\mathcal{Z}_j : (6.40) \text{ holds}\},$$
$$\mathcal{Z}_j^\ddagger = \{(m_2,m_3)\in\mathcal{Z}_j : (6.43) \text{ holds}\}.$$

Step 6. Using (6.37) and taking into account that the identity (5.36) holds whenever $\mathbf{m}^*=(m_2,m_3)\neq(0,0)$, we now have the trivial upper bound

$$\int_0^1\int_0^1 \Xi(\alpha_1,\alpha_2)\,\mathrm{d}\alpha_1\,\mathrm{d}\alpha_2 \leqslant \mathfrak{I}_1+\mathfrak{I}_2, \tag{6.44}$$

where

$$\mathfrak{I}_1 = \sum_{j=0}^{\infty}\sum_{\ell=0}^{\infty}\sum_{\mathbf{m}^*\in\mathcal{Z}_j^\dagger}\int_{\Omega_j(\mathbf{m}^*;\ell)}\Lambda(\mathcal{C}(f;u);H;\mathbf{v};\mathbf{m})\,\mathrm{d}\alpha_1\,\mathrm{d}\alpha_2, \tag{6.45}$$

$$\mathfrak{I}_2 = \sum_{j=0}^{\infty}\sum_{\ell=0}^{\infty}\sum_{\mathbf{m}^*\in\mathcal{Z}_j^\ddagger}\int_{\Omega_j(\mathbf{m}^*;\ell)}\Lambda(\mathcal{C}(f;u);H;\mathbf{v};\mathbf{m})\,\mathrm{d}\alpha_1\,\mathrm{d}\alpha_2. \tag{6.46}$$

Lemma 6.4. *We have*

$$\mathfrak{J}_1 \leqslant \frac{c_{19}}{u^{1/2}}, \tag{6.47}$$

where the constant $c_{19} = c_{19}(f; a, b) > 0$ depends at most on the function f in the interval $[a, b]$.

Proof. Suppose that $(m_2, m_3) \in \mathcal{Z}_j^\dagger$. It follows from (5.38) and (6.41) that

$$c_{20} 2^j \leqslant \min\{|m_2|, |m_3|\} \leqslant \max\{|m_2|, |m_3|\} \leqslant 2^j, \tag{6.48}$$

where the constant $c_{20} = c_{20}(f; a, b) > 0$ satisfies $2c_{18}c_{20} = 1$.

If $(\alpha_1, \alpha_2) \in \Omega_j(\mathbf{m}^*; 0)$, then it follows from (5.35), (6.36), (6.38) and (6.39) that

$$\Lambda(\mathcal{C}(f; u); H; \mathbf{v}; \mathbf{m}) \leqslant \frac{c_{17} H}{|m_3|^{1/2}} \min\left\{u, \frac{1}{u(\pi m_3)^2}\right\}. \tag{6.49}$$

Using (5.41), (6.48) and (6.49), we deduce that

$$\begin{aligned}
&\sum_{\mathbf{m}^* \in \mathcal{Z}_j^\dagger} \int_{\Omega_j(\mathbf{m}^*; 0)} \Lambda(\mathcal{C}(f; u); H; \mathbf{v}; \mathbf{m}) \, \mathrm{d}\alpha_1 \, \mathrm{d}\alpha_2 \\
&\quad \leqslant 4 \sum_{c_{20} 2^j \leqslant m_2 \leqslant 2^j} \sum_{c_{20} 2^j \leqslant m_3 \leqslant 2^j} \lambda_2(\Omega_j(\mathbf{m}^*; 0)) \Lambda(\mathcal{C}(f; u); H; \mathbf{v}; \mathbf{m}) \\
&\quad \leqslant 2^8 c_{17} \sum_{c_{20} 2^j \leqslant m_3 \leqslant 2^j} \frac{2^j}{m_3^{1/2}} \min\left\{u, \frac{1}{u(\pi m_3)^2}\right\}.
\end{aligned} \tag{6.50}$$

Note next that

$$\sum_{c_{20} 2^j \leqslant m_3 \leqslant 2^j} \frac{2^j}{m_3^{1/2}} \min\left\{u, \frac{1}{u(\pi m_3)^2}\right\} \leqslant \sum_{c_{20} 2^j \leqslant m_3 \leqslant 2^j} \frac{2^j}{u\pi^2 m_3^{5/2}} \leqslant \frac{2^{2j}}{u\pi^2 (c_{20} 2^j)^{5/2}}. \tag{6.51}$$

Combining (6.50) and (6.51), we deduce that

$$\sum_{\mathbf{m}^* \in \mathcal{Z}_j^\dagger} \int_{\Omega_j(\mathbf{m}^*; 0)} \Lambda(\mathcal{C}(f; u); H; \mathbf{v}; \mathbf{m}) \, \mathrm{d}\alpha_1 \, \mathrm{d}\alpha_2 \leqslant \frac{c_{21}}{u 2^{j/2}}, \tag{6.52}$$

where the constant $c_{21} = c_{21}(f; a, b) > 0$ satisfies $c_{21} c_{20}^{5/2} \pi^2 = 2^8 c_{17}$. Note also that

$$\sum_{c_{20} 2^j \leqslant m_3 \leqslant 2^j} \frac{2^j}{m_3^{1/2}} \min\left\{u, \frac{1}{u(\pi m_3)^2}\right\} \leqslant \frac{u 2^{2j}}{(c_{20} 2^j)^{1/2}} = \frac{u 2^{3j/2}}{c_{20}^{1/2}}. \tag{6.53}$$

Combining (6.50) and (6.53), we deduce that

$$\sum_{\mathbf{m}^* \in \mathcal{Z}_j^\dagger} \int_{\Omega_j(\mathbf{m}^*; 0)} \Lambda(\mathcal{C}(f; u); H; \mathbf{v}; \mathbf{m}) \, \mathrm{d}\alpha_1 \, \mathrm{d}\alpha_2 \leqslant c_{22} u 2^{3j/2}, \tag{6.54}$$

where the constant $c_{22} = c_{22}(f; a, b) > 0$ satisfies $c_{22} c_{20}^{1/2} = 2^8 c_{17}$.

If $(\alpha_1, \alpha_2) \in \Omega_j(\mathbf{m}^*; \ell)$, where $\ell \geqslant 1$, then it follows from (5.35), (5.40), (6.36), (6.38) and (6.39) that

$$\Lambda(\mathcal{C}(f; u); H; \mathbf{v}; \mathbf{m}) \leqslant \frac{2^4 c_{17} H}{\pi^2 4^\ell |m_3|^{1/2}} \min\left\{u, \frac{1}{u(\pi m_3)^2}\right\}. \tag{6.55}$$

Using (5.41), (6.48), (6.51) and (6.55), we deduce that

$$\begin{aligned}
&\sum_{\mathbf{m}^*\in\mathcal{Z}_j^\dagger}\int_{\Omega_j(\mathbf{m}^*;\ell)}\Lambda(\mathcal{C}(f;u);H;\mathbf{v};\mathbf{m})\,\mathrm{d}\alpha_1\,\mathrm{d}\alpha_2\\
&\quad\leqslant 4\sum_{c_{20}2^j\leqslant m_2\leqslant 2^j}\ \sum_{c_{20}2^j\leqslant m_3\leqslant 2^j}\lambda_2(\Omega_j(\mathbf{m}^*;\ell))\Lambda(\mathcal{C}(f;u);H;\mathbf{v};\mathbf{m})\\
&\quad\leqslant\frac{2^{12}c_{17}}{\pi^2 2^\ell}\sum_{c_{20}2^j\leqslant m_3\leqslant 2^j}\frac{2^j}{m_3^{1/2}}\min\left\{u,\frac{1}{u(\pi m_3)^2}\right\}\leqslant\frac{c_{23}}{u2^\ell 2^{j/2}},
\end{aligned}$$

where the constant $c_{23}=c_{23}(f;a,b)>0$ depends at most on the function f in the interval $[a,b]$, so that

$$\sum_{\ell=1}^{\infty}\sum_{\mathbf{m}^*\in\mathcal{Z}_j^\dagger}\int_{\Omega_j(\mathbf{m}^*;\ell)}\Lambda(\mathcal{C}(f;u);H;\mathbf{v};\mathbf{m})\,\mathrm{d}\alpha_1\,\mathrm{d}\alpha_2\leqslant\frac{c_{23}}{u2^{j/2}}. \tag{6.56}$$

Meanwhile, using (6.53) instead of (6.51), we deduce that

$$\sum_{\mathbf{m}^*\in\mathcal{Z}_j^\dagger}\int_{\Omega_j(\mathbf{m}^*;\ell)}\Lambda(\mathcal{C}(f;u);H;\mathbf{v};\mathbf{m})\,\mathrm{d}\alpha_1\,\mathrm{d}\alpha_2\leqslant\frac{c_{24}u2^{3j/2}}{2^\ell},$$

where the constant $c_{24}=c_{24}(f;a,b)>0$ depends at most on the function f in the interval $[a,b]$, so that

$$\sum_{\ell=1}^{\infty}\sum_{\mathbf{m}^*\in\mathcal{Z}_j^\dagger}\int_{\Omega_j(\mathbf{m}^*;\ell)}\Lambda(\mathcal{C}(f;u);H;\mathbf{v};\mathbf{m})\,\mathrm{d}\alpha_1\,\mathrm{d}\alpha_2\leqslant c_{24}u2^{3j/2}. \tag{6.57}$$

Let $J^\dagger$ denote the largest non-negative integer such that

$$2^j\leqslant\frac{1}{u},$$

so that

$$2^{J^\dagger}\leqslant\frac{1}{u}\quad\text{and}\quad 2^{J^\dagger+1}>\frac{1}{u}. \tag{6.58}$$

Combining (6.45), (6.52), (6.54), (6.56), and (6.57), we conclude that

$$\mathfrak{I}_1\leqslant c_{25}u\sum_{j=0}^{J^\dagger}2^{3j/2}+\frac{c_{26}}{u}\sum_{j=J^\dagger+1}^{\infty}\frac{1}{2^{j/2}}, \tag{6.59}$$

where the constant $c_{25}=c_{25}(f;a,b)>0$ satisfies $c_{25}=\max\{c_{22},c_{24}\}$ while the constant $c_{26}=c_{26}(f;a,b)>0$ satisfies $c_{26}=\max\{c_{21},c_{23}\}$. The inequality (6.47) follows on combining (6.58) and (6.59). $\square$

Lemma 6.5. *We have*

$$\mathfrak{I}_2\leqslant c_{27}, \tag{6.60}$$

where the constant $c_{27}=c_{27}(f;a,b)>0$ depends at most on the function f in the interval $[a,b]$.

Proof. The proof is in three parts.

Part 1. Here we obtain new bounds for the term $\widetilde{\mathcal{I}}_1(\mathbf{m};C(f;u))$ in the case when the inequality (6.43) holds. Indeed, combining (6.2) and (6.43), we obtain the lower bound

$$\min_{a\leqslant x\leqslant b}|m_2+m_3f'(x)|\geqslant\frac{1}{4}\min\{c_6,1\}\max\{|m_2|,|m_3|\}. \tag{6.61}$$

We make use the simple identity

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\mathrm{e}^{2\pi\mathrm{i}(m_2x+m_3f(x))}}{2\pi\mathrm{i}(m_2+m_3f'(x))}\right)=\mathrm{e}^{2\pi\mathrm{i}(m_2x+m_3f(x))}+\frac{\mathrm{i}m_3\mathrm{e}^{2\pi\mathrm{i}(m_2x+m_3f(x))}f''(x)}{2\pi(m_2+m_3f'(x))^2}. \tag{6.62}$$

Combining (6.62) and integration by parts, we obtain

$$\begin{aligned}
&\int_{a+\gamma_1}^{b-\gamma_1}\mathrm{e}^{2\pi\mathrm{i}(m_2x+m_3f(x))}f'(x)\,\mathrm{d}x+\mathrm{i}\int_{a+\gamma_1}^{b-\gamma_1}\frac{m_3\mathrm{e}^{2\pi\mathrm{i}(m_2x+m_3f(x))}f''(x)}{2\pi(m_2+m_3f'(x))^2}f'(x)\,\mathrm{d}x\\
&\quad=\int_{a+\gamma_1}^{b-\gamma_1}\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\mathrm{e}^{2\pi\mathrm{i}(m_2x+m_3f(x))}}{2\pi\mathrm{i}(m_2+m_3f'(x))}\right)f'(x)\,\mathrm{d}x\\
&\quad=\left[\frac{\mathrm{e}^{2\pi\mathrm{i}(m_2x+m_3f(x))}}{2\pi\mathrm{i}(m_2+m_3f'(x))}f'(x)\right]_{a+\gamma_1}^{b-\gamma_1}+\mathrm{i}\int_{a+\gamma_1}^{b-\gamma_1}\frac{\mathrm{e}^{2\pi\mathrm{i}(m_2x+m_3f(x))}}{2\pi(m_2+m_3f'(x))}f''(x)\,\mathrm{d}x. \qquad (6.63)
\end{aligned}$$

Combining (6.32) and (6.63), we have

$$\widetilde{\mathcal{I}}_1(\mathbf{m};C(f;u))=\mathrm{i}\,\widetilde{\mathcal{I}}_1^{(1)}-\mathrm{i}\,\widetilde{\mathcal{I}}_1^{(2)}+\mathrm{i}\,\widetilde{\mathcal{I}}_1^{(3)}-\mathrm{i}\,\widetilde{\mathcal{I}}_1^{(4)}, \tag{6.64}$$

where

$$\widetilde{\mathcal{I}}_1^{(1)}=\frac{2}{b-a}\int_0^{(b-a)/2}\int_{a+\gamma_1}^{b-\gamma_1}\frac{\mathrm{e}^{2\pi\mathrm{i}(m_2x+m_3f(x))}}{2\pi(m_2+m_3f'(x))}f''(x)\,\mathrm{d}x\,\mathrm{d}\gamma_1, \tag{6.65}$$

$$\widetilde{\mathcal{I}}_1^{(2)}=\frac{2}{b-a}\int_0^{(b-a)/2}\int_{a+\gamma_1}^{b-\gamma_1}\frac{m_3\mathrm{e}^{2\pi\mathrm{i}(m_2x+m_3f(x))}f''(x)}{2\pi(m_2+m_3f'(x))^2}f'(x)\,\mathrm{d}x\,\mathrm{d}\gamma_1, \tag{6.66}$$

$$\begin{aligned}
\widetilde{\mathcal{I}}_1^{(3)}&=\frac{2}{b-a}\int_0^{(b-a)/2}\frac{\mathrm{e}^{2\pi\mathrm{i}(m_2(a+\gamma_1)+m_3f(a+\gamma_1))}}{2\pi(m_2+m_3f'(a+\gamma_1))}f'(a+\gamma_1)\,\mathrm{d}\gamma_1\\
&=\frac{2}{b-a}\int_a^{(a+b)/2}\frac{\mathrm{e}^{2\pi\mathrm{i}(m_2(x)+m_3f(x))}}{2\pi(m_2+m_3f'(x))}f'(x)\,\mathrm{d}x, \qquad (6.67)
\end{aligned}$$

$$\begin{aligned}
\widetilde{\mathcal{I}}_1^{(4)}&=\frac{2}{b-a}\int_0^{(b-a)/2}\frac{\mathrm{e}^{2\pi\mathrm{i}(m_2(b-\gamma_1)+m_3f(b-\gamma_1))}}{2\pi(m_2+m_3f'(b-\gamma_1))}f'(b-\gamma_1)\,\mathrm{d}\gamma_1\\
&=\frac{2}{b-a}\int_{(a+b)/2}^{b}\frac{\mathrm{e}^{2\pi\mathrm{i}(m_2(x)+m_3f(x))}}{2\pi(m_2+m_3f'(x))}f'(x)\,\mathrm{d}x. \qquad (6.68)
\end{aligned}$$

To study the integral $\widetilde{\mathcal{I}}_1^{(1)}$, note that

$$\int_{a+\gamma_1}^{b-\gamma_1}\frac{\mathrm{e}^{2\pi\mathrm{i}(m_2x+m_3f(x))}}{2\pi(m_2+m_3f'(x))}f''(x)\,\mathrm{d}x=\int_{a+\gamma_1}^{b-\gamma_1}\mathrm{e}^{\mathrm{i}F(x)}G(x)\,\mathrm{d}x,$$

where

$$F(x)=2\pi(m_2x+m_3f(x))\quad\text{and}\quad G(x)=\frac{f''(x)}{2\pi(m_2+m_3f'(x))}, \tag{6.69}$$

so that

$$\frac{F'(x)}{G(x)}=\frac{4\pi^2(m_2+m_3f'(x))^2}{f''(x)}=4\pi^2H(x),$$

where

$$H(x)=\frac{(m_2+m_3f'(x))^2}{f''(x)}.$$

Using (6.3), (6.61) and (6.69), we deduce the lower bound

$$\left|\frac{F'(x)}{G(x)}\right|=\frac{4\pi^2|m_2+m_3f'(x)|^2}{|f''(x)|}\geqslant c_{28}(\max\{|m_2|,|m_3|\})^2,$$

where the constant $c_{28} = c_{28}(f; a, b) > 0$ depends at most on the function f in the interval $[a, b]$. Furthermore, the technical condition implies that we can apply Lemma 6.2 to each of the at most c_{12} subintervals of $[a + \gamma_1, b - \gamma_1]$, and this leads to the bound

$$\left|\int_{a+\gamma_1}^{b-\gamma_1} \frac{\mathrm{e}^{2\pi(m_2x+m_3f(x))}}{2\pi\mathrm{i}(m_2+m_3f'(x))} f''(x)\,\mathrm{d}x\right| \leqslant \frac{c_{29}}{(\max\{|m_2|,|m_3|\})^2},$$

where the constant $c_{29} = c_{29}(f; a, b) > 0$ depends at most on the function f in the interval $[a, b]$. It then follows trivially from (6.65) that

$$|\widetilde{\mathcal{I}}_1^{(1)}| \leqslant \frac{c_{29}}{(\max\{|m_2|,|m_3|\})^2}. \tag{6.70}$$

To study the integral $\widetilde{\mathcal{I}}_1^{(2)}$, note that

$$\int_{a+\gamma_1}^{b-\gamma_1} \frac{m_3\mathrm{e}^{2\pi\mathrm{i}(m_2x+m_3f(x))}f''(x)}{2\pi(m_2+m_3f'(x))^2} f'(x)\,\mathrm{d}x = \int_{a+\gamma_1}^{b-\gamma_1} \mathrm{e}^{\mathrm{i}F(x)}G(x)\,\mathrm{d}x,$$

where

$$F(x) = 2\pi(m_2x + m_3f(x)) \quad\text{and}\quad G(x) = \frac{m_3f'(x)f''(x)}{2\pi(m_2+m_3f'(x))^2}, \tag{6.71}$$

so that

$$\frac{F'(x)}{G(x)} = \frac{4\pi^2(m_2+m_3f'(x))^3}{m_3f'(x)f''(x)} = \frac{4\pi^2}{m_3}H(x),$$

where

$$H(x) = \frac{(m_2+m_3f'(x))^3}{f'(x)f''(x)}.$$

Using (6.3), (6.61) and (6.71), we deduce the lower bound

$$\left|\frac{F'(x)}{G(x)}\right| = \frac{4\pi^2|m_2+m_3f'(x)|^3}{|m_3|\,|f'(x)|\,|f''(x)|} \geqslant c_{30}(\max\{|m_2|,|m_3|\})^2$$

where the constant $c_{30} = c_{30}(f; a, b) > 0$ depends at most on the function f in the interval $[a, b]$. Furthermore, the technical condition implies that we can apply Lemma 6.2 to each of the at most c_{12} subintervals of $[a + \gamma_1, b - \gamma_1]$, and this leads to the bound

$$\left|\int_{a+\gamma_1}^{b-\gamma_1} \frac{m_3\mathrm{e}^{2\pi\mathrm{i}(m_2x+m_3f(x))}f''(x)}{2\pi(m_2+m_3f'(x))^2} f'(x)\,\mathrm{d}x\right| \leqslant \frac{c_{31}}{(\max\{|m_2|,|m_3|\})^2},$$

where the constant $c_{31} = c_{31}(f; a, b) > 0$ depends at most on the function f in the interval $[a, b]$. It then follows trivially from (6.66) that

$$|\widetilde{\mathcal{I}}_1^{(2)}| \leqslant \frac{c_{31}}{(\max\{|m_2|,|m_3|\})^2}. \tag{6.72}$$

To study the integrals $\widetilde{\mathcal{I}}_1^{(3)}$ and $\widetilde{\mathcal{I}}_1^{(4)}$ given by (6.67) and (6.68), note that

$$\int_a^{(a+b)/2} \frac{\mathrm{e}^{2\pi\mathrm{i}(m_2(x)+m_3f(x))}}{2\pi(m_2+m_3f'(x))} f'(x)\,\mathrm{d}x = \int_a^{(a+b)/2} \mathrm{e}^{\mathrm{i}F(x)}G(x)\,\mathrm{d}x,$$

$$\int_{(a+b)/2}^{b} \frac{\mathrm{e}^{2\pi\mathrm{i}(m_2(x)+m_3f(x))}}{2\pi(m_2+m_3f'(x))} f'(x)\,\mathrm{d}x = \int_{(a+b)/2}^{b} \mathrm{e}^{\mathrm{i}F(x)}G(x)\,\mathrm{d}x,$$

where

$$F(x) = 2\pi(m_2x + m_3f(x)) \quad\text{and}\quad G(x) = \frac{f'(x)}{2\pi(m_2+m_3f'(x))}, \tag{6.73}$$

so that

$$\frac{F'(x)}{G(x)} = \frac{4\pi^2(m_2 + m_3 f'(x))^2}{f'(x)} = 4\pi^2 H(x),$$

where

$$H(x) = \frac{(m_2 + m_3 f'(x))^2}{f'(x)}.$$

Using (6.3), (6.61) and (6.73), we deduce the lower bound

$$\left|\frac{F'(x)}{G(x)}\right| = \frac{4\pi^2|m_2 + m_3 f'(x)|^2}{|f'(x)|} \geqslant c_{32}(\max\{|m_2|, |m_3|\})^2,$$

where the constant $c_{32} = c_{32}(f; a, b) > 0$ depends at most on the function f in the interval $[a, b]$. Furthermore, the derivative $H'(x)$ is equal to

$$\frac{2m_3 f'(x) f''(x)(m_2 + m_3 f'(x)) - f''(x)(m_2 + m_3 f'(x))^2}{(f'(x))^2},$$

where the denominator is non-zero and the numerator is equal to

$$f''(x)(m_2 + m_3 f'(x))(m_3 f'(x) - m_2).$$

Here, in view of (6.3), the factor $f''(x)$ is non-zero, and the factors $m_2 + m_3 f'(x)$ and $m_3 f'(x) - m_2$ can each have at most one zero in $[a, b]$. Hence the interval $[a, b]$ is a union of at most 3 subintervals such that $F'(x)/G(x)$ is monotonic in each subinterval. We can apply Lemma 6.2 to each of these subintervals, and this leads to the bounds

$$|\widetilde{\mathcal{I}}_1^{(3)}| \leqslant \frac{c_{33}}{(\max\{|m_2|, |m_3|\})^2} \quad \text{and} \quad |\widetilde{\mathcal{I}}_1^{(4)}| \leqslant \frac{c_{33}}{(\max\{|m_2|, |m_3|\})^2}, \tag{6.74}$$

where the constant $c_{33} = c_{33}(f; a, b) > 0$ depends at most on the function f in the interval $[a, b]$.

Combining (6.64), (6.70), (6.72) and (6.74), we obtain the bound

$$|\widetilde{\mathcal{I}}_1(\mathbf{m}; C(f; u))| \leqslant \frac{c_{34}}{(\max\{|m_2|, |m_3|\})^2}, \tag{6.75}$$

where the constant $c_{34} = c_{34}(f; a, b) > 0$ satisfies $c_{34} = 4\max\{c_{29}, c_{31}, c_{33}\}$.

Part 2. The idea is to combine the bound (6.75) for the term $\widetilde{\mathcal{I}}_1(\mathbf{m}; C(f; u))$ with our earlier bounds (6.39) for $\widetilde{\mathcal{I}}_2(\mathbf{m}; C(f; u))$ and (5.35) for $\widetilde{\mathcal{J}}(\mathbf{m}; \mathbf{v}; H)$. However, we need some care and make some refinement to the sets $\mathcal{Z}_j$, $j = 0, 1, 2, 3, \ldots$, discussed in Step 4 in Section 5 and defined by (5.38).

For every $j = 0, 1, 2, 3, \ldots$ and $k = 0, 1, \ldots, j$, let

$$\mathcal{Z}_{j,k} = \begin{cases} \{(m_2, m_3) \in \mathcal{Z}_j : 2^{k-1} < \min\{|m_2|, |m_3|\} \leqslant 2^k\}, & \text{if } k \neq 0, \\ \{(m_2, m_3) \in \mathcal{Z}_j : 0 \leqslant \min\{|m_2|, |m_3|\} \leqslant 1\}, & \text{if } k = 0, \end{cases} \tag{6.76}$$

so that

$$|\mathcal{Z}_{j,k}| \leqslant 2^{j+k+3}, \tag{6.77}$$

and the set $\mathcal{Z}_j$ can be written as a disjoint union

$$\mathcal{Z}_j = \bigcup_{k=0}^{j} \mathcal{Z}_{j,k}.$$

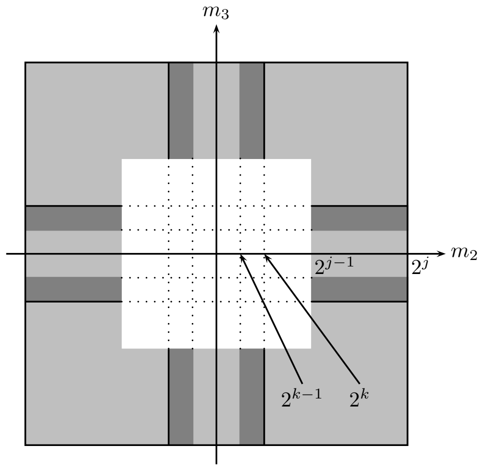

Figure 6.2: the location of $\mathbf{m}^* = (m_2, m_3) \in \mathcal{Z}_{j,k}$

It then follows from (6.46) that

$$\mathfrak{I}_2 = \sum_{j=0}^{\infty} \sum_{k=0}^{j} \sum_{\ell=0}^{\infty} \sum_{\mathbf{m}^* \in \mathcal{Z}_{j,k}} \int_{\Omega_j(\mathbf{m}^*;\ell)} \Lambda(\mathcal{C}(f;u); H; \mathbf{v}; \mathbf{m}) \, \mathrm{d}\alpha_1 \, \mathrm{d}\alpha_2. \tag{6.78}$$

Part 3. Suppose that $(m_2, m_3) \in \mathcal{Z}_{j,k}$, where $k \neq 0$. It follows from (5.38) and (6.76) that

$$2^{k-1} < \min\{|m_2|, |m_3|\} < 2^k \quad \text{and} \quad 2^{j-1} < \max\{|m_2|, |m_3|\} < 2^j. \tag{6.79}$$

Suppose furthermore that $(\alpha_1, \alpha_2) \in \Omega_j(\mathbf{m}^*; 0)$. Then it follows from (5.35), (6.36), (6.39), (6.75) and (6.79) that

$$\begin{aligned}
\Lambda(\mathcal{C}(f;u); H; \mathbf{v}; \mathbf{m}) &\leqslant \frac{c_{34} H}{(\max\{|m_2|, |m_3|\})^2} \min\left\{u, \frac{1}{u(\pi m_3)^2}\right\} \\
&\leqslant \frac{c_{34} H}{(\max\{|m_2|, |m_3|\})^2} \min\left\{u, \frac{1}{u\pi^2 (\min\{|m_2|, |m_3|\})^2}\right\} \\
&\leqslant \frac{c_{34} H}{2^{2j-2}} \min\left\{u, \frac{1}{u\pi^2 2^{2k-2}}\right\}.
\end{aligned}$$

Combining this with (5.41) and (6.77), we deduce that

$$\begin{aligned}
&\sum_{\mathbf{m}^* \in \mathcal{Z}_{j,k}} \int_{\Omega_j(\mathbf{m}^*;0)} \Lambda(\mathcal{C}(f;u); H; \mathbf{v}; \mathbf{m}) \, \mathrm{d}\alpha_1 \, \mathrm{d}\alpha_2 \\
&\qquad \leqslant \frac{2^{j+k+9}}{H} \frac{c_{34} H}{2^{2j-2}} \min\left\{u, \frac{1}{u\pi^2 2^{2k-2}}\right\} = \frac{2^{11} c_{34}}{2^j} \min\left\{2^k u, \frac{4}{u\pi^2 2^k}\right\}.
\end{aligned} \tag{6.80}$$

It is easily checked that the inequality (6.80) holds for $k = 0$ also.

Let $K^\ddagger$ denote the largest integer k such that

$$2^k u \leqslant \frac{4}{u\pi^2 2^k}, \quad \text{equivalent to} \quad 2^k \leqslant \frac{2}{u\pi}.$$

Then

$$2^{K^\ddagger} \leqslant \frac{2}{u\pi} < \frac{1}{u} \quad \text{and} \quad 2^{K^\ddagger+1} > \frac{2}{u\pi}, \tag{6.81}$$

and

$$\begin{aligned}\sum_{k=0}^{j}\min\left\{2^k u,\frac{4}{u\pi^2 2^k}\right\} &\leqslant \sum_{k=0}^{K^\ddagger}2^k u+\frac{4}{u\pi^2}\left(\frac{1}{2^{K^\ddagger+1}}+\ldots+\frac{1}{2^j}\right)\\ &\leqslant 2^{K^\ddagger+1}u+\frac{8}{u\pi^2 2^{K^\ddagger+1}}<4, \qquad (6.82)\end{aligned}$$

in view of (6.81). It now follows from (6.80) and (6.82) that

$$\sum_{j=0}^{\infty}\sum_{k=0}^{j}\sum_{\mathbf{m}^*\in\mathcal{Z}_{j,k}}\int_{\Omega_j(\mathbf{m}^*;0)}\Lambda(\mathcal{C}(f;u);H;\mathbf{v};\mathbf{m})\,\mathrm{d}\alpha_1\,\mathrm{d}\alpha_2\leqslant\sum_{j=0}^{\infty}\frac{2^{13}c_{34}}{2^j}\leqslant c_{35}, \qquad (6.83)$$

where the constant $c_{35}=c_{35}(f;a,b)>0$ depends at most on the function f in the interval $[a,b]$.

Suppose next that $(m_2,m_3)\in\mathcal{Z}_{j,k}$, where $k\neq 0$, and $(\alpha_1,\alpha_2)\in\Omega_j(\mathbf{m}^*;\ell)$, where $\ell\geqslant 1$. Then it follows from (5.35), (5.40), (6.36), (6.39), (6.75) and (6.79) that

$$\begin{aligned}\Lambda(\mathcal{C}(f;u);H;\mathbf{v};\mathbf{m}) &\leqslant \frac{2c_{34}}{(\max\{|m_2|,|m_3|\})^2}\min\left\{u,\frac{1}{u(\pi m_3)^2}\right\}\frac{2}{\pi^2 H(\mathbf{v}\cdot\mathbf{m})^2}\\ &\leqslant \frac{2c_{34}H}{2^{2\ell}(\max\{|m_2|,|m_3|\})^2}\min\left\{u,\frac{1}{u\pi^2(\min\{|m_2|,|m_3|\})^2}\right\}\\ &\leqslant \frac{2c_{34}H}{2^{2\ell}2^{2j-2}}\min\left\{u,\frac{1}{u\pi^2 2^{2k-2}}\right\}.\end{aligned}$$

Combining this with (5.41) and (6.77), we deduce that

$$\begin{aligned}&\sum_{\mathbf{m}^*\in\mathcal{Z}_{j,k}}\int_{\Omega_j(\mathbf{m}^*;\ell)}\Lambda(\mathcal{C}(f;u);H;\mathbf{v};\mathbf{m})\,\mathrm{d}\alpha_1\,\mathrm{d}\alpha_2\\ &\quad\leqslant\frac{2^{j+k+10}2^\ell}{H}\frac{c_{34}H}{2^{2\ell}2^{2j-2}}\min\left\{u,\frac{1}{u\pi^2 2^{2k-2}}\right\}=\frac{2^{12}c_{34}}{2^\ell 2^j}\min\left\{2^k u,\frac{4}{u\pi^2 2^k}\right\}. \qquad (6.84)\end{aligned}$$

It now follows from (6.82) and (6.84) that

$$\begin{aligned}&\sum_{j=0}^{\infty}\sum_{k=0}^{j}\sum_{\ell=1}^{\infty}\sum_{\mathbf{m}^*\in\mathcal{Z}_{j,k}}\int_{\Omega_j(\mathbf{m}^*;\ell)}\Lambda(\mathcal{C}(f;u);H;\mathbf{v};\mathbf{m})\,\mathrm{d}\alpha_1\,\mathrm{d}\alpha_2\\ &\quad\leqslant\sum_{j=0}^{\infty}\sum_{k=0}^{j}\sum_{\ell=1}^{\infty}\frac{2^{12}c_{34}}{2^\ell 2^j}\min\left\{2^k u,\frac{4}{u\pi^2 2^k}\right\}\leqslant\sum_{j=0}^{\infty}\frac{2^{14}c_{34}}{2^j}\leqslant 2c_{35}. \qquad (6.85)\end{aligned}$$

The desired inequality (6.60) follows on combining (6.78), (6.83) and (6.85). □

It now follows from (6.44), (6.47) and (6.60) that

$$\int_0^1\int_0^1\Xi(\alpha_1,\alpha_2)\,\mathrm{d}\alpha_1\,\mathrm{d}\alpha_2\leqslant\psi(u),$$

where $\psi(u)$ is given by (6.6). Then for every real parameter $\kappa>1$, we have

$$\lambda_2(\{(\alpha_1,\alpha_2)\in[0,1]^2:\Xi(\alpha_1,\alpha_2)\geqslant\kappa\psi(u)\})\leqslant\frac{1}{\kappa}.$$

Step 7. We can now justify all the steps in a similar way as in Step 7 in Section 5.

This completes the proof of Lemma 6.1. □

References

[1] J. Beck, W.W.L. Chen, Y. Yang. New Kronecker–Weyl type equidistribution results and diophantine approximation. *Eur. J. Math.* **9** (2023), article 38 (87 pp.).

[2] J. Beck, W.W.L. Chen, Y. Yang. Uniformity of geodesic flow in non-integrable 3-manifolds (manuscript).

[3] M. Drmota, R.F. Tichy. *Sequences, Discrepancies and Applications* (Lecture Notes in Mathematics **1651**, Springer, 1997).

[4] D. König, A. Szücs. Mouvement d'un point abondonne a l'interieur d'un cube. *Rend. Circ. Mat. Palermo* **36** (1913), 79–90.

[5] E.C. Titchmarsh. *The Theory of the Riemann Zeta-function*, 2nd edition revised by D.R. Heath-Brown (Oxford Science Publications, Clarendon Press, 1986).

[6] W.A. Veech. Strict ergodicity in zero dimensional dynamical systems and the Kronecker–Weyl theorem mod 2. *Trans. Amer. Math. Soc.* **140** (1969), 1–33.

Department of Mathematics, Hill Center for the Mathematical Sciences, Rutgers University, Piscataway NJ 08854, USA
Email address: jbeck@math.rutgers.edu

School of Mathematical and Physical Sciences, Faculty of Science and Engineering, Macquarie University, Sydney NSW 2109, Australia
Email address: william.chen@mq.edu.au

School of Science, Beijing University of Posts and Telecommunications, Beijing 100876, China
Email address: yangyx@bupt.edu.cn