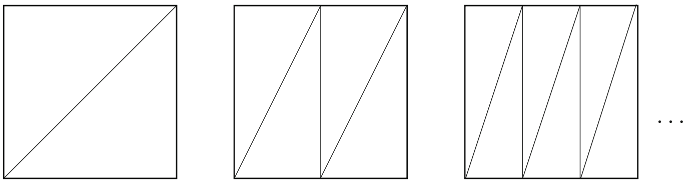
One square and an odd number of triangles

Chapter 22

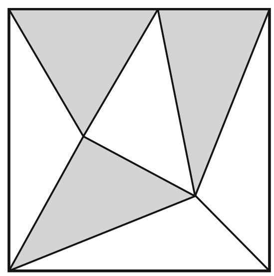
Suppose we want to dissect a square into n triangles of equal area. When n is even, this is easily accomplished. For example, you could divide the horizontal sides into $\frac{n}{2}$ segments of equal length and draw a diagonal in each of the $\frac{n}{2}$ rectangles:



But now assume n is odd. Already for $n = 3$ this causes problems, and after some experimentation you will probably come to think that it might not be possible. So let us pose the general problem:

> *Is it possible to dissect a square into an odd number n of triangles of equal area?*

Now, this looks like a classical question of Euclidean geometry, and one could have guessed that surely the answer must have been known for a long time (if not to the Greeks). But when Fred Richman and John Thomas popularized the problem in the 1960s they found to their surprise that no one knew the answer or a reference where this would be discussed.



There are dissections of squares into an odd number of triangles whose areas are *nearly* equal.

Well, the answer is "no" not only for $n = 3$, but for any odd n. But how should one prove a result like this? By scaling we may, of course, restrict ourselves to the unit square with vertices $(0,0), (1,0), (0,1), (1,1)$. Any argument must therefore somehow make use of the fact that the area of the triangles in a dissection is $\frac{1}{n}$, where n is odd. The following proof due to Paul Monsky, with initial work of John Thomas, is a stroke of genius and totally unexpected: It uses an algebraic tool, valuations, to construct a striking coloring of the plane, and combines this with some elegant and stunningly simple combinatorial reasonings. And what's more: at present no other proof is known!

Before we state the theorem let us prepare the ground by a quick study of valuations. Everybody is familiar with the absolute value function $|x|$ on the rationals $\mathbb{Q}$ (or the reals $\mathbb{R}$). It maps $\mathbb{Q}$ to the nonnegative reals such that for all x and y,

M. Aigner, G. M. Ziegler, *Proofs from THE BOOK*, https://doi.org/10.1007/978-3-662-57265-8_22

(i) $|x| = 0$ if and only if $x = 0$,

(ii) $|xy| = |x||y|$, and

(iii) $|x + y| \leq |x| + |y|$ (the triangle inequality).

The triangle inequality makes $\mathbb{R}$ into a metric space and gives rise to the familiar notions of convergence. It was a great discovery around 1900 that besides the absolute value there are other natural "value functions" on $\mathbb{Q}$ that satisfy the conditions (i) to (iii).

Let p be a prime number. Any rational number $r \neq 0$ can be written uniquely in the form

$$r = p^k \frac{a}{b}, \quad k \in \mathbb{Z}, \tag{1}$$

where a and $b > 0$ are relatively prime to p. Define the *p-adic value*

$$|r|_p := p^{-k}, \quad |0|_p = 0. \tag{2}$$

Example: $|\frac{3}{4}|_2 = 4$, $|\frac{6}{7}|_2 = |2|_2 = \frac{1}{2}$, and $|\frac{3}{4} + \frac{6}{7}|_2 = |\frac{45}{28}|_2 = |\frac{1}{4} \cdot \frac{45}{7}|_2 = 4 = \max\{|\frac{3}{4}|_2, |\frac{6}{7}|_2\}$.

Conditions (i) and (ii) are obviously satisfied, and for (iii) we obtain the even stronger inequality

(iii′) $|x + y|_p \leq \max\{|x|_p, |y|_p\}$ (the non-Archimedean property).

Indeed, let $r = p^k \frac{a}{b}$ and $s = p^\ell \frac{c}{d}$, where we may assume that $k \geq \ell$, that is, $|r|_p = p^{-k} \leq p^{-\ell} = |s|_p$. Then we get

$$\begin{aligned} |r+s|_p &= \Big|p^k \frac{a}{b} + p^\ell \frac{c}{d}\Big|_p = \Big|p^\ell (p^{k-\ell} \frac{a}{b} + \frac{c}{d})\Big|_p \\ &= p^{-\ell} \Big|\frac{p^{k-\ell}ad + bc}{bd}\Big|_p \leq p^{-\ell} = \max\{|r|_p, |s|_p\}, \end{aligned}$$

since the denominator bd is relatively prime to p. We also see from this that

(iv) $|x + y|_p = \max\{|x|_p, |y|_p\}$ whenever $|x|_p \neq |y|_p$,

but we will prove below that this property is quite generally implied by (iii′).

Any function $v : K \to \mathbb{R}_{\geq 0}$ on a field K that satisfies

(i) $v(x) = 0$ if and only if $x = 0$,

(ii) $v(xy) = v(x)v(y)$, and

(iii′) $v(x + y) \leq \max\{v(x), v(y)\}$ (non-Archimedean property)

for all $x, y \in K$ is called a *non-Archimedean real valuation* of K.

For every such valuation v we have $v(1) = v(1)v(1)$, hence $v(1) = 1$; and $1 = v(1) = v((-1)(-1)) = [v(-1)]^2$, so $v(-1) = 1$. Thus from (ii) we get $v(-x) = v(x)$ for all x and $v(x^{-1}) = v(x)^{-1}$ for $x \neq 0$.

Every field has the *trivial* valuation that maps every nonzero element onto 1, and if v is a real non-Archimedean valuation, then so is v^t for any positive real number t. So for $\mathbb{Q}$ we have the p-adic valuations and their powers, and a famous theorem of Ostrowski states that any nontrivial real non-Archimedean valuation of $\mathbb{Q}$ is of this form.

As announced, let us verify that the important property

(iv) $v(x+y) \;=\; \max\{v(x), v(y)\}$ if $v(x) \neq v(y)$

holds for any non-Archimedean valuation. Indeed, suppose that we have $v(x) < v(y)$. Then

$$\begin{aligned} v(y) \;=\; v((x+y)-x) &\;\leq\; \max\{v(x+y), v(x)\} \;=\; v(x+y) \\ &\;\leq\; \max\{v(x), v(y)\} \;=\; v(y) \end{aligned}$$

where (iii′) yields the inequalities, the first equality is clear, and the other two follow from $v(x) < v(y)$. Thus $v(x+y) = v(y) = \max\{v(x), v(y)\}$.

The property (iv) together with $v(-x) = v(x)$ also implies that $v(a \pm b_1 \pm b_2 \pm \cdots \pm b_\ell) = v(a)$ if $v(a) > v(b_i)$ for all i.

Monsky's beautiful approach to the square dissection problem used an extension of the 2-adic valuation $|x|_2$ to a valuation v of $\mathbb{R}$, where "extension" means that we require $v(x) = |x|_2$ whenever x is in $\mathbb{Q}$. Such a non-Archimedean real extension exists, but this is not standard algebra fare. In the following, we present Monsky's argument in a version due to Hendrik Lenstra that requires much less; it only needs a valuation v that takes values in an arbitrary "ordered group", not necessarily in $(\mathbb{R}_{>0}, \cdot, <)$, such that $v(\frac{1}{2}) > 1$. The definition and the existence of such a valuation will be provided in the appendix to this chapter.

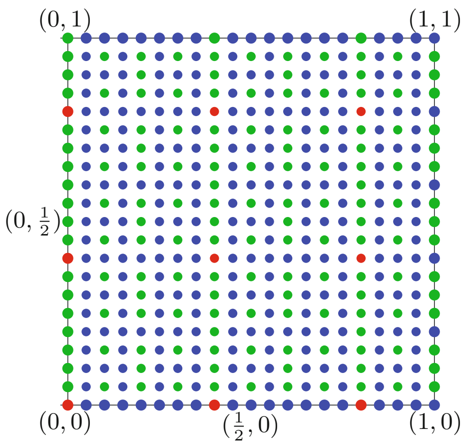
Here we just note that any valuation with $v(\frac{1}{2}) > 1$ satisfies $v(\frac{1}{n}) = 1$ for odd integers n. Indeed, $v(\frac{1}{2}) > 1$ means that $v(2) < 1$, and thus $v(2k) < 1$ by (iii′) and induction on k. From this we get $v(2k+1) = 1$ from (iv), and thus again $v(\frac{1}{2k+1}) = 1$ from (ii).

Monsky's Theorem. *It is not possible to dissect a square into an odd number of triangles of equal area.*

■ **Proof.** In the following we construct a specific three-coloring of the plane with amazing properties. One of them is that the area of any triangle whose vertices have three different colors — which in the following is called a *rainbow triangle* — has a v-value larger than 1, so the area cannot be $\frac{1}{n}$ for odd n. And then we verify that any dissection of the unit square must contain such a rainbow triangle, and the proof will be complete.

The coloring of the points (x, y) of the real plane will be constructed by looking at the entries of the triple $(x, y, 1)$ that have the maximal value under the valuation v. This maximum may occur once or twice or even three times. The color (blue, or green, or red) will record the coordinate of $(x, y, 1)$ in which the maximal v-value occurs first:

$$(x,y) \text{ is colored } \begin{cases} \text{blue} & \text{if } v(x) \geq v(y),\ v(x) \geq v(1), \\ \text{green} & \text{if } v(x) < v(y),\ v(y) \geq v(1), \\ \text{red} & \text{if } v(x) < v(1),\ v(y) < v(1). \end{cases}$$



This assigns a unique color to each point in the plane. The figure in the margin shows the color for each point in the unit square whose coordinates are fractions of the form $\frac{k}{20}$.

The following statement is the first step to the proof.

Lemma 1. *For any blue point $p_b = (x_b, y_b)$, green point $p_g = (x_g, y_g)$, and red point $p_r = (x_r, y_r)$, the v-value of the determinant*

$$\det\begin{pmatrix} x_b & y_b & 1 \\ x_g & y_g & 1 \\ x_r & y_r & 1 \end{pmatrix}$$

is at least 1.

■ **Proof.** The determinant is a sum of six terms. One of them is the product of the entries of the main diagonal, $x_b y_g 1$. By construction of the coloring each of the diagonal entries compared to the other entries in the row has a maximal v-value, so comparing with the last entry in each row (which is 1) we get

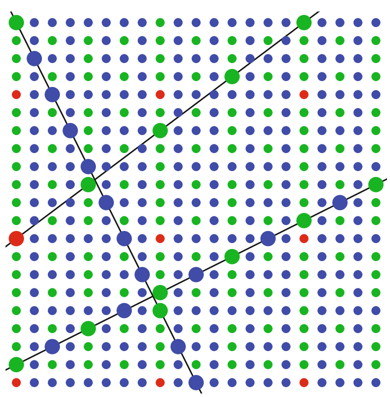
$$v(x_b y_g 1) \;=\; v(x_b)v(y_g)v(1) \;\geq\; v(1)v(1)v(1) \;=\; 1.$$

Any of the other five summands of the determinant is a product of three matrix entries, one from each row (with a sign that as we know is irrelevant for the v-value). It picks at least one matrix entry below the main diagonal, whose v-value is strictly smaller than that of the diagonal entry in the same row, and at least one matrix entry above the main diagonal, whose v-value is not larger than that of the diagonal entry in the same row. Thus all of the five other summands of the determinant have a v-value that is strictly smaller than the summand corresponding to the main diagonal. Thus by property (iv) of non-Archimedean valuations, we find that the v-value of the determinant is given by the summand corresponding to the main diagonal,

$$v\Big(\det\begin{pmatrix} x_b & y_b & 1 \\ x_g & y_g & 1 \\ x_r & y_r & 1 \end{pmatrix}\Big) \;=\; v(x_b y_g 1) \;\geq\; 1.$$

□

Corollary. *Any line of the plane receives at most two different colors. The area of a rainbow triangle cannot be* 0*, and it cannot be $\frac{1}{n}$ for odd n.*



■ **Proof.** The area of the triangle with vertices at a blue point p_b, a green point p_g, and a red point p_r is the absolute value of

$$\tfrac{1}{2}\big((x_b - x_r)(y_g - y_r) - (x_g - x_r)(y_b - y_r)\big),$$

which up to the sign is half the determinant of Lemma 1.

The three points cannot lie on a line since the determinant cannot be 0, as $v(0) = 0$. The area of the triangle cannot be $\frac{1}{n}$, since in this case we would get $\pm\frac{2}{n}$ for the determinant, thus

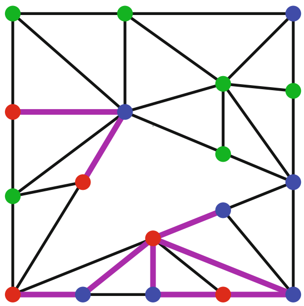
$$v(\pm\tfrac{2}{n}) \;=\; v(\tfrac{1}{2})^{-1}v(\tfrac{1}{n}) \;<\; 1$$

because of $v(\frac{1}{2}) > 1$ and $v(\frac{1}{n}) = 1$, contradicting Lemma 1. □

And why did we construct this coloring? Because we are now going to show that in *any* dissection of the unit square $S = [0,1]^2$ into triangles (equal-sized or not!) there must always be a rainbow triangle, which according to the corollary cannot have area $\frac{1}{n}$ for odd n. Thus the following lemma will complete the proof of Monsky's theorem.

Lemma 2. *Every dissection of the unit square $S = [0,1]^2$ into finitely many triangles contains an odd number of rainbow triangles, and thus at least one.*

■ **Proof.** The following counting argument is truly inspired. The idea is due to Emanuel Sperner, and will reappear with "Sperner's Lemma" in Chapter 28.



Consider the segments between neighboring vertices in a given dissection. A segment is called a *red-blue segment* if one endpoint is red and the other is blue. For the example in the figure, the red-blue segments are drawn in purple.

We make two observations, repeatedly using the fact from the corollary that on any line there can be points of at most two colors.

(A) The bottom line of the square contains an *odd* number of red-blue segments, since $(0,0)$ is red and $(1,0)$ is blue, and all vertices in between are red or blue. So on the walk from the red end to the blue end of the bottom line, there must be an odd number of changes between red and blue. The other boundary lines of the square contain no red-blue segments.

(B) If a triangle T has at most two colors at its vertices, then it contains an *even* number of red-blue segments on its boundary. However, every rainbow triangle has an *odd* number of red-blue segments on its boundary.

Indeed, there is an odd number of red-blue segments between a red vertex and a blue vertex of a triangle, but an even number (if any) between any vertices with a different color combination. Thus a rainbow triangle has an odd number of red-blue segments in its boundary, while any other triangle has an even number (two or zero) of vertex pairs with the color combination red and blue.

Now let us count the boundary red-blue segments summed over all triangles in the dissection. Since every red-blue segment in the interior of the square is counted twice, and there is an odd number on the boundary of S, this count is *odd*. Hence we conclude from **(B)** that there must be an odd number of rainbow triangles. □

Appendix: Extending valuations

It is not at all obvious that an extension of a non-Archimedean real valuation from one field to a larger one is always possible. But it can be done, not only from $\mathbb{Q}$ to $\mathbb{R}$, but generally from any field K to a field L that contains K. (This is known as "Chevalley's theorem"; see for example the book by Jacobson [1].)

In the following, we establish much less — but enough for our application to odd dissections. Indeed, in our proof for Monsky's theorem we have not used the addition for values of $v : \mathbb{R} \to \mathbb{R}_{\geq 0}$; we have used only the multiplication and the order on $\mathbb{R}_{\geq 0}$. Hence for our argument it is sufficient if the nonzero values of v lie in a (multiplicatively written) *ordered abelian group* $(G, \cdot, <)$. That is, the elements of G are linearly ordered, and $a < b$ in G implies $ac < bc$ for any $a, b, c \in G$. As we assume that the group is written multiplicatively, the neutral element of G is denoted by 1. For the definition of a valuation, we adjoin a special element 0 with the understanding that $0 \notin G$, $0a = 0$, and $0 < a$ hold for all $a \in G$. Of course, the prime example of an ordered abelian group is $(\mathbb{R}_{>0}, \cdot, <)$ with the usual linear order, and the prime example for $\{0\} \cup G$ is $(\mathbb{R}_{\geq 0}, \cdot)$.

Definition. Let K be a field. A *non-Archimedean valuation* v with values in an ordered abelian group G is a map $v : K \to \{0\} \cup G$ with

(i) $v(x) = 0 \Longleftrightarrow x = 0$,

(ii) $v(xy) = v(x)v(y)$,

(iii′) $v(x + y) \leq \max\{v(x), v(y)\}$, and

(iv) $v(x + y) = \max\{v(x), v(y)\}$ whenever $v(x) \neq v(y)$

for all $x, y \in K$.

The fourth condition in this description is again implied by the first three. And among the simple consequences we record that if $v(x) < 1$, $x \neq 0$, then $v(x^{-1}) = v(x)^{-1} > 1$.

So here is what we will establish:

Theorem. *The field of real numbers $\mathbb{R}$ has a non-Archimedean valuation to an ordered abelian group*

$$v : \mathbb{R} \to \{0\} \cup G$$

such that $v(\frac{1}{2}) > 1$.

■ **Proof.** We first relate any valuation on a field to a subring of the field. (All the subrings that we consider contain 1.) Suppose $v : K \to \{0\} \cup G$ is a valuation; let

$$R := \{x \in K : v(x) \leq 1\}, \qquad U := \{x \in K : v(x) = 1\}.$$

It is immediate that R is a subring of K, called the *valuation ring* corresponding to v. Furthermore, $v(xx^{-1}) = v(1) = 1$ implies that $v(x) = 1$

if and only if $v(x^{-1}) = 1$. Thus U is the set of units (invertible elements) of R. In particular, U is a subgroup of $K^\times$, where we write $K^\times := K \setminus \{0\}$ for the multiplicative group of K. Finally, with $R^{-1} := \{x^{-1} : x \neq 0\}$ we have $K = R \cup R^{-1}$. Indeed, if $x \notin R$ then $v(x) > 1$ and therefore $v(x^{-1}) < 1$, thus $x^{-1} \in R$. The property $K = R \cup R^{-1}$ already characterizes all possible valuation rings in a given field.

Lemma. *A proper subring $R \subseteq K$ is a valuation ring with respect to some valuation v into some ordered group G if and only if $K = R \cup R^{-1}$.*

■ **Proof.** We have seen one direction. Suppose now $K = R \cup R^{-1}$. How should we construct the group G? If $v : K \to \{0\} \cup G$ is a valuation corresponding to R, then $v(x) < v(y)$ holds if and only if $v(xy^{-1}) < 1$, that is, if and only if $xy^{-1} \in R \setminus U$. Also, $v(x) = v(y)$ if and only if $xy^{-1} \in U$, or $xU = yU$ as cosets in the factor group $K^\times/U$.

Hence the natural way to proceed goes as follows. Take the quotient group $G := K^\times/U$, and define an order relation on G by setting

$$xU < yU \quad :\Longleftrightarrow \quad xy^{-1} \in R \setminus U.$$

It is a nice exercise to check that this indeed makes G into an ordered group. The map $v : K \to \{0\} \cup G$ is then defined in the most natural way:

$$v(0) := 0, \qquad \text{and} \qquad v(x) := xU \ \text{ for } x \neq 0.$$

It is easy to verify conditions (i) to (iii′) for v, and that R is the valuation ring corresponding to v. □

In order to establish the theorem, it thus suffices to find a valuation ring $B \subseteq \mathbb{R}$ such that $\frac{1}{2} \notin B$.

Claim. *Any inclusion-maximal subring $B \subseteq \mathbb{R}$ with the property $\frac{1}{2} \notin B$ is a valuation ring.*

First we should perhaps note that a *maximal* subring $B \subseteq \mathbb{R}$ with the property $\frac{1}{2} \notin B$ exists. This is not quite trivial — but it does follow with a routine application of Zorn's lemma, which is reviewed in the box. Indeed, if we have an ascending chain of subrings $B_i \subseteq \mathbb{R}$ that don't contain $\frac{1}{2}$, then this chain has an upper bound, given by the union of all the subrings B_i, which again is a subring and does not contain $\frac{1}{2}$.

$\mathbb{Z} \subseteq \mathbb{R}$ is such a subring with $\frac{1}{2} \notin \mathbb{Z}$, but it is not maximal.

Zorn's Lemma

The Lemma of Zorn is of fundamental importance in algebra and other parts of mathematics when one wants to construct maximal structures. It also plays a decisive role in the logical foundations of mathematics.

Lemma. *Suppose $P_\leq$ is a nonempty partially ordered set with the property that every ascending chain $(a_i)_\leq$ has an upper bound b, such that $a_i \leq b$ for all i. Then $P_\leq$ contains a maximal element M, meaning that there is no $c \in P$ with $M < c$.*

To prove the Claim, let us assume that $B \subseteq \mathbb{R}$ is a maximal subring not containing $\frac{1}{2}$. If B is not a valuation ring, then there is some element $\alpha \in \mathbb{R}\backslash(B \cup B^{-1})$. We denote by $B[\alpha]$ the subring generated by $B \cup \alpha$, that is, the set of all real numbers that can be written as polynomials in α with coefficients in B. Let $2B \subseteq B$ be the subset of all elements of the form $2b$, for $b \in B$. Now $2B$ is a subset of B, so we have $2B[\alpha] \subseteq B[\alpha]$ and $2B[\alpha^{-1}] \subseteq B[\alpha^{-1}]$. If we had $2B[\alpha] \neq B[\alpha]$ or $2B[\alpha^{-1}] \neq B[\alpha^{-1}]$, then due to $1 \in B$ this would imply that $\frac{1}{2} \notin B[\alpha]$ resp. $\frac{1}{2} \notin B[\alpha^{-1}]$, contradicting the maximality of $B \subseteq \mathbb{R}$ as a subring that does not contain $\frac{1}{2}$. Thus we get that $2B[\alpha] = B[\alpha]$ and $2B[\alpha^{-1}] = B[\alpha^{-1}]$. This implies that $1 \in B$ can be written in the form

$$1 = 2u_0 + 2u_1\alpha + \cdots + 2u_m\alpha^m \quad \text{with } u_i \in B, \quad (1)$$

and similarly as

$$1 = 2v_0 + 2v_1\alpha^{-1} + \cdots + 2v_n\alpha^{-n} \text{ with } v_i \in B, \quad (2)$$

which after multiplication by α^n and subtraction of $2v_0\alpha^n$ from both sides yields

$$(1 - 2v_0)\alpha^n = 2v_1\alpha^{n-1} + \cdots + 2v_{n-1}\alpha + 2v_n. \quad (3)$$

Let us assume that these representations are chosen such that m and n are as small as possible. We may also assume that $m \geq n$, otherwise we exchange α with α^{-1}, and (1) with (2).

Now multiply (1) by $1 - 2v_0$ and add $2v_0$ on both sides of the equation, to get

$$1 = 2(u_0(1 - 2v_0) + v_0) + 2u_1(1 - 2v_0)\alpha + \cdots + 2u_m(1 - 2v_0)\alpha^m.$$

But if in this equation we substitute for the term $(1-2v_0)\alpha^m$ the expression given by equation (3) multiplied by α^{m-n}, then this results in an equation that expresses $1 \in B$ as a polynomial in $2B[\alpha]$ of degree at most $m - 1$. This contradiction to the minimality of m establishes the Claim. □

References

[1] N. JACOBSON: *Lectures in Abstract Algebra, Part III: Theory of Fields and Galois Theory,* Graduate Texts in Mathematics 32, Springer, New York 1975.

[2] P. MONSKY: *On dividing a square into triangles,* Amer. Math. Monthly **77** (1970), 161-164.

[3] F. RICHMAN & J. THOMAS: *Problem 5471,* Amer. Math. Monthly **74** (1967), 329.

[4] S. K. STEIN & S. SZABÓ *:* Algebra and Tiling: Homomorphisms in the Service of Geometry, Carus Math. Monographs **25**, MAA, Washington DC 1994.

[5] J. THOMAS: *A dissection problem,* Math. Magazine **41** (1968), 187-190.