

MATH 47A PROBLEM SETS, FALL 2005

DMITRY KLEINBOCK

ABSTRACT. The first three sections contain all the homework assignments written in a research paper format. In the last section I wrote down a solution of one of the problems to illustrate the theorem-proof environment.

1. INVARIANTS AND SEMI-INVARIANTS

Reading material: [E, pp. 1–23].

Due date: September 21, 2005.

1.1. Four grasshoppers sit at four vertices of a square. Every second one of the grasshoppers jumps over another one and lands at the symmetric point (that is, if a grasshopper jumps from point A over B and lands at C , then vectors \overrightarrow{AB} and \overrightarrow{BC} are equal).

Prove that:

- (a) no three of them will ever sit on a straight line parallel to one of the sides of the original square (more generally – on any straight line);
- (b) the four of them will never form a square bigger than the original one (more generally – a parallelogram with area bigger than the area of the original one).

1.2. The set of numbers $\{a, b, c\}$ every second gets replaced with the set $\{a+b-c, b+c-a, c+a-b\}$. In the beginning $a = 2004$, $b = 2005$, $c = 2006$. Can we get $\{2004, 2006, 2007\}$ after a sequence of these operations?

1.3. On a board the 100 numbers $1, \frac{1}{2}, \frac{1}{3}, \dots, \frac{1}{100}$ are written down. Every second we choose two numbers a and b from those on the board, erase them and write the number $a+b+ab$. This operation is performed 99 times until there is just one number. What is this number? Find it and prove that it does not depend on the sequence of choices.

Date: September 27, 2005.

1.4. Initially there is a pile of 637 chips on a table. We are allowed to throw away one chip and divide the pile into two (not necessarily equal) piles. Then the same can be done with any pile containing more than two chips, and so on. Is it possible to get only the piles consisting of three chips?

1.5. Stones are arranged in three piles: in one – 51 stones, in another one – 49 stones, and in the third one – 5 stones. It is allowed to combine any two piles into one, and also to divide a pile with even number of stones into two equal piles. Is it possible to obtain 105 piles with a single stone in each?

1.6. On a circle there are several blue points and several red points. It is allowed to add a red point and change colors of its two neighbors, and also, if there are more than two points, to remove a red point and change colors of its former neighbors. Suppose that initially there were 2 red points and no blue points. Prove that it is impossible to obtain a configuration consisting of:

- (a) (easy) 3 blue points and no red points;
- (b) (difficult!) 2 blue points and no red points.

2. COLORING PROOFS

Reading material: [E, pp. 25–37].

Due date: October 3, 2005.

2.1. 9 chips occupy the lower left 3×3 square of an 8×8 chessboard. For one move any chip can jump over any other (not necessarily adjacent) chip and land into the cell symmetric to the first cell around the second one. It is possible after a number of moves to arrange all the chips in the form of the upper right 3×3 square?

2.2. A piece of cheese has the form of a $3 \times 3 \times 3$ cube from which the central cube is cut out. A mouse starts to gnaw this piece of cheese. First she eats some $1 \times 1 \times 1$ cube. After the mouse finishes up each little $1 \times 1 \times 1$ cube, she goes on to one of the adjacent (on a side) cubes. Can she eat all the cheese?

2.3. A king, moving according to the rules of chess, visited all the cells of the standard chessboard exactly once and returned to its original location. Prove that it made an even number of diagonal moves.

2.4. A convex n -gon is divided into triangles by non-intersecting diagonals so that an odd number of triangles meet in each of its vertices. Prove that n is divisible by 3.

2.5. 99 squares of size 2×2 are cut from a 29×29 square sheet of graph paper. Prove that it is possible to cut out another 2×2 square from the remaining part of the big square.

2.6. The points on the boundary of a regular triangle are colored in two colors. Prove that among them there exist three points of the same color forming a right triangle.

3. THE EXTREMAL PRINCIPLE

Reading material: [E, pp. 39–57].

Due date: October 12, 2005.

3.1. In a circle, a finite set S of chords (segments connecting two points on the circle) is given, with a property that each of the chords from S passes through a midpoint of another chords from S . Prove that all these chords are diameters (that is, connect the antipodal points).

[**Hint:** consider the chord of minimal length.]

3.2. Prove that it is not possible to find different natural numbers x, y, z, t which are solutions of

$$x^x + y^y = z^z + t^t. \quad (3.1)$$

[**Hint:** look at the maximal number among x, y, z, t .]

3.3. Each of the $3n$ members of a parliament slapped one of his/her colleagues. Prove that among them it is possible to choose a committee consisting of n members none of whom slapped each other.

[**Hint:** consider those lucky ones who got slapped the least number of times.]

3.4. 30 numbers a_1, \dots, a_{30} are written along the circle in such a way that each of them is equal to the absolute value of the difference between the next two in the clockwise direction (that is, $a_1 = |a_2 - a_3|$, $a_2 = |a_3 - a_4|$, \dots , $a_{29} = |a_{30} - a_1|$, $a_{30} = |a_1 - a_2|$). The sum of all of the numbers is equal to 20. What are they?

[**Hint:** start by looking at the largest number.]

3.5. Natural numbers are placed at each of the 8 vertices of a cube in such a way that numbers at adjacent vertices (i.e. those sharing an edge) differ by no more than 1. Prove that one can find two opposite vertices of the cube such that the numbers placed there also differ by no more than 1.

[**Hint:** consider a vertex containing the smallest number.]

4. THE BOX PRINCIPLE

Reading material: [E, pp. 59–83].

Due date: October 19, 2005.

4.1. Each of 102 students in a school is a friend of at least 68 others. Prove that among them one can choose four who have the same number of friends. (Assume that friendship is a symmetric relation.)

4.2. 2003 different prime numbers form an arithmetic progression, and the smallest of the numbers is greater than 2003. Prove that the common difference of the progression is divisible by 2003.

4.3. Suppose that a quadratic polynomial $ax^2 + bx + c$ is equal to the fourth power of an integer for all positive integers x . Prove that $a = b = 0$.

4.4. 2000 real numbers are written in a row. Prove that it is possible to choose a subset consisting of adjacent numbers (maybe of just one number) such that the sum of numbers from this subset differs from an integer by no more than $1/1000$.

4.5. An open subset U of $[0, 1]$ has the property that for any $x, y \in U$, the distance between x and y is different from $1/10$. Prove that the measure (length) of U is not greater than $1/2$.

4.6. 50 segments are given on a straight line. Suppose that no point on the line belongs to more than 7 of the segments. Prove that one can find 8 of them which are pairwise disjoint.

5. A SOLUTION TO A SAMPLE PROBLEM

Here I want to pick one of the problems, say Problem 1.2 from §1, and write down its solution as if it were a mathematical paper, illustrating some features of L^AT_EX.

First let me make a theorem out of the statement of the problem.

Theorem 5.1. *Let T be a map from \mathbb{R}^3 to itself defined by*

$$T((a, b, c)) \stackrel{\text{def}}{=} (a + b - c, b + c - a, c + a - b). \quad (5.1)$$

Then $\forall n \in \mathbb{N}$,

$$T^n((2004, 2005, 2006)) \neq (2004, 2006, 2007). \quad (5.2)$$

And here is a solution in the form of a rigorous mathematical proof of Theorem 5.1.

Proof. Define a function $F : \mathbb{R}^3 \rightarrow \mathbb{R}$ by

$$F((a, b, c)) \stackrel{\text{def}}{=} a + b + c. \quad (5.3)$$

Then, using (5.1), for any $(a, b, c) \in \mathbb{R}^3$ one can write

$$\begin{aligned} F\left(T((a, b, c))\right) &= f((a + b - c, b + c - a, c + a - b)) = \\ &= a + b - c + b + c - a + c + a - b = a + b + c = F((a, b, c)). \end{aligned}$$

Arguing by induction, one deduces that

$$F\left(T^n((a, b, c))\right) = F((a, b, c)) \quad \forall (a, b, c) \in \mathbb{R}^3 \text{ and } \forall n \in \mathbb{N}. \quad (5.4)$$

However¹ $F((2004, 2005, 2006))$ is not equal to $F((2004, 2006, 2007))$. Thus (5.2) follows from (5.4). \square

¹This can be established by a direct computation

6. (C, α) -GOOD FUNCTIONS

Homework problems due: Monday, November 14

Let V be a subset of \mathbb{R} and f a function on V . In what follows, we will let $\|f\|_B \stackrel{\text{def}}{=} \sup_{x \in B} |f(x)|$ for a subset B of V . Lebesgue measure of B will be denoted by $|B|$.

For positive numbers C and α , say that f is (C, α) -good on V if for any open interval $B \subset V$ one has

$$\forall \varepsilon > 0 \quad |\{x \in B : |f(x)| < \varepsilon\}| \leq C \left(\frac{\varepsilon}{\|f\|_B} \right)^\alpha |B|. \quad (6.1)$$

Here we adopt the convention $\frac{1}{0} = \infty$, so that (6.1) holds if $f|_B \equiv 0$.

The properties listed below follow immediately from the definition.

Lemma 6.1. *Let $V \subset \mathbb{R}$ and $C, \alpha > 0$ be given.*

- (a) f is (C, α) -good on $V \iff$ so is $|f|$;
- (b) f is (C, α) -good on $V \implies$ so is $cf \ \forall c \in \mathbb{R}$;
- (c) If f is (C, α) -good on V and $c_1 \leq \frac{|f(x)|}{|g(x)|} \leq c_2$ for all $x \in V$, then g is $(C(c_2/c_1)^\alpha, \alpha)$ -good on V ;
- (d) f is (C, α) -good on $V \implies$ it is (C', α') -good on V' for every $C' \geq C$, $\alpha' \leq \alpha$ and $V' \subset V$.
- (e) $f_i, i \in I$, are (C, α) -good on $V \implies$ so is $\sup_{i \in I} |f_i|$.

Problem 6.1. Prove (e).

The notion of (C, α) -good functions was introduced in [KM] in 1998, but the importance of (6.1) for certain estimates arising in flows on homogeneous spaces was observed earlier. For instance, the next proposition, which describes what can be called a model example of good functions, can already be found in a 1993 paper by Dani and Margulis.

Proposition 6.2. [DM, Lemma 4.1] *For any $n \in \mathbb{N}$, any polynomial $f \in \mathbb{R}[x]$ of degree not greater than n is $(C_n, 1/n)$ -good on \mathbb{R} , where $C_n = 2n(n+1)^{1/n}$.*

Proof. We are going to prove an even stronger statement, with $2n$ above replaced by n . Fix an open interval $B \subset \mathbb{R}$, a polynomial $f \in \mathbb{R}[x]$ of degree not exceeding n , and a positive ε . We need to show that

$$|\{x \in B : |f(x)| < \varepsilon\}| \leq n(n+1)^{1/n} \left(\frac{\varepsilon}{\|f\|_B} \right)^{1/n} |B|. \quad (6.2)$$

Claim 6.3. *Let $A \subset \mathbb{R}$ be such that $|A| > m$. Then there exist $x_1, \dots, x_{n+1} \in A$ with $|x_i - x_j| \geq m/n$ for each $1 \leq i \neq j \leq n+1$.*

Problem 6.2. Prove Claim 6.3.

Now assume that the left hand side of (6.2) is strictly bigger than some number m , and choose $x_1, \dots, x_{n+1} \in B$ as in Claim 6.3 applied to $A = \{x \in B : |f(x)| < \varepsilon\}$. Using Lagrange's interpolation formula one can write down the exact expression for f :

$$f(x) = \sum_{i=1}^{n+1} f(x_i) \frac{\prod_{j=1, j \neq i}^{n+1} (x - x_j)}{\prod_{j=1, j \neq i}^{n+1} (x_i - x_j)}. \quad (6.3)$$

Note that $|f(x_i)| < \varepsilon$ for each i , $|x - x_j| \leq |B|$ for each j and $x \in B$, and also $|x_i - x_j| \geq m/n$. Therefore

$$\|f\|_B < (n+1)\varepsilon \frac{|B|^n}{(m/n)^n}.$$

which can be rewritten as

$$m < n(n+1)^{1/n} \left(\frac{\varepsilon}{\|f\|_B} \right)^{1/n} |B|,$$

proving (6.2). □

Problem 6.3. Improve Proposition 6.2 by observing that the estimate $|x_i - x_j| \geq m/n$ can be strengthened to $|x_i - x_j| \geq m|i - j|/n$. Try to obtain the best possible value of C_n . What is its asymptotics as $n \rightarrow \infty$?

Problem 6.4. Find a lower bound for C_n by considering 'worst possible' polynomials. Hopefully this will prove that the upper bound obtained in the previous problem is optimal.

The next two problems are optional (so far), and strictly speaking are not problems but rather invitations to think further about these topics.

Problem 6.5. A 'big picture' question: define functions which are (C, α) -good on a subset of \mathbb{R}^2 , and try to prove that polynomials of two variables are (C, α) -good on \mathbb{R}^2 for some C, α . What are C and α ? What is a two-dimensional analogue of Claim 6.3? how does Lagrange interpolation work in two variables? or maybe it is better to use something else for the proof?

Problem 6.6. An even bigger question: what if instead of polynomials we take functions with uniform estimates on their n th derivatives? for simplicity start with $n = 2$, fix $\delta > 0$, and consider twice continuously differentiable functions f on a segment V such that for some $c > 0$ one has

$$c \leq |f''(x)| \leq c(1 + \delta) \quad \text{for all } x \in V.$$

Can you prove that they are all (C, α) -good on V with C dependent only on δ (preferably with $\alpha = 1/2$)? This is clearly a problem harder than the one about polynomials (for example, the set of functions with the above property is an infinite-dimensional space). Does Lagrange interpolation work? I mean, certainly not precisely, but maybe it is possible to estimate the value of f from above if one knows that it takes small values at three points far away from each other? (Note: this is NOT how we attacked this problem in [KM].) Is it possible to obtain a result which will tend to Proposition 6.2 as $\delta \rightarrow 0$?

7. (C, α) -GOOD FUNCTIONS OF SEVERAL VARIABLES

Homework problems due: Monday, December 12

A follow-up on Problem 6.5: the notion of good functions can be easily extended to functions of several real variables and, more generally, functions on arbitrary metric spaces. Namely let X be a metric space and $|\cdot|$ a measure on X which assigns positive values to all nonempty open sets (i.e. has full support) and is finite on all bounded sets. Let V be a subset of X and f a real-valued function on V . For $C, \alpha > 0$, say that f is (C, α) -good on V if (6.1) holds for any open ball $B \subset V$. Properties listed in Lemma 6.1 clearly hold in this generality.

The metric space we will be interested in is \mathbb{R}^d . To exploit its product structure it will be convenient to use the “product metric” in which balls in the product space are products of the balls in the factors. That is, for vectors $\mathbf{x} = (x_1, \dots, x_d)$ and $\mathbf{y} = (y_1, \dots, y_d)$ in \mathbb{R}^d , we will work with $\text{dist}(\mathbf{x}, \mathbf{y}) \stackrel{\text{def}}{=} \max_i |x_i - y_i|$, as opposed to the Euclidean distance $\text{dist}_E(\mathbf{x}, \mathbf{y}) \stackrel{\text{def}}{=} \sqrt{|x_1 - y_1|^2 + \dots + |x_d - y_d|^2}$. The next problem explains why it does not matter much which distance to use.

Problem 7.1. Show that there exists a constant C_d depending only on d such that: whenever f is (C, α) -good on \mathbb{R}^d with respect to the metric “dist”, it is $(C_d C, \alpha)$ -good on \mathbb{R}^d with respect to the metric “dist_E”, and vice versa.

For the next theorem we take two metric spaces X and Y with measures of full support both denoted by $|\cdot|$, and the metric on $X \times Y$ is the product metric, $\text{dist}((x, y), (x', y')) = \max(\text{dist}_X(x, x'), \text{dist}_Y(y, y'))$.

Theorem 7.1. [KT] *Suppose f is a function on $U \times V$, where $U \subset X$ and $V \subset Y$ are open subsets, and suppose C, D, α, β are positive constants such that*

$$\text{for all } y \in V, \text{ the function } x \mapsto f(x, y) \text{ is } (C, \alpha)\text{-good on } U, \quad (7.1)$$

and

for all $x \in U$, the function $y \mapsto f(x, y)$ is (D, β) -good on V , (7.2)

Then f is (E, γ) -good on $U \times V$, where

$$\gamma = \frac{\alpha\beta}{\alpha + \beta} \quad \text{and} \quad E = (\alpha + \beta) \left(\left(\frac{C}{\beta} \right)^\beta \left(\frac{D}{\alpha} \right)^\alpha \right)^{\frac{1}{\alpha + \beta}}. \quad (7.3)$$

Proof. Fix a ball in $U \times V$ of the form $A \times B$, where A and B are balls in X and Y . Without loss of generality let us rescale the measures and f so that $|A| = |B| = \|f\|_{\mu \times \nu, A \times B} = 1$.

Problem 7.2. Explain why it is possible to do this.

To simplify the exposition, let us introduce the following notation: for a ball B in a metric space, a function f on B and a positive ε , define

$$B^{f, \varepsilon} \stackrel{\text{def}}{=} \{x \in B : |f(x)| < \varepsilon\}.$$

Take an arbitrary $\varepsilon > 0$; we need to demonstrate that

$$|(A \times B)^{f, \varepsilon}| \leq E\varepsilon^\gamma. \quad (7.4)$$

For $y \in B$ let us denote by f_y the function $x \mapsto f(x, y)$. Also denote by φ the function defined on B by $\varphi(y) \stackrel{\text{def}}{=} \|f_y\|_A$; note that $\|\varphi\|_B = 1$ because of the assumed normalization of f . In view of (7.1), for any $y \in B$ one has

$$|A^{f_y, \varepsilon}| \leq C \left(\frac{\varepsilon}{\varphi(y)} \right)^\alpha \iff \varphi(y) \leq \left(\frac{C}{|A^{f_y, \varepsilon}|} \right)^{1/\alpha} \varepsilon. \quad (7.5)$$

Then take an arbitrary $t > 0$ (to be fixed later), and denote

$$B_t \stackrel{\text{def}}{=} \{y \in B : |A^{f_y, \varepsilon}| \geq t\}.$$

In view of (7.5), $y \in B_t$ implies that $\varphi(y)$ is not bigger than $(C/t)^{1/\alpha} \varepsilon$. Since it follows from Lemma 6.1 and (7.2) that φ is (D, β) -good on V , one can write

$$|B_t| \leq |(\{y \in B : \varphi(y) \leq (C/t)^{1/\alpha} \varepsilon\})| \leq D \left((C/t)^{1/\alpha} \varepsilon \right)^\beta. \quad (7.6)$$

Now observe that one has $|(\{x \in A : (x, y) \in (A \times B)^{f, \varepsilon}\})| < t$ whenever $y \notin B_t$, therefore, by the Fubini Theorem,

$$\begin{aligned} |(A \times B)^{f, \varepsilon}| &< |A \times B_t| + t \cdot |B \setminus B_t| \\ &\leq |B_t| + t \stackrel{(7.6)}{\leq} t + (DC^{\beta/\alpha} \varepsilon^\beta) \cdot t^{-\beta/\alpha}. \end{aligned} \quad (7.7)$$

At this point let us recall that we have not chosen t yet. Since the function in the right hand side of (7.7) consists of two summands, one increasing and the other one decreasing, it seems natural to choose the value of t at which they are equal.

Problem 7.3. Show that choosing the value of t as above and plugging it into (7.7) gives γ as in (7.3) and $E = 2(C^\beta D^\alpha)^{\frac{1}{\alpha+\beta}}$.

However this does not produce the best constant:

Problem 7.4. Show that the right hand side of (7.7) attains its global minimum when $t = \left(C^\beta \left(\frac{D\beta}{\alpha}\right)^\alpha\right)^{\frac{1}{\alpha+\beta}} \varepsilon^{\frac{\alpha\beta}{\alpha+\beta}}$, and that substituting it into (7.7) gives E and γ as in (7.3).

This finishes the proof of the theorem. \square

Problem 7.5. Using induction, prove the following

Corollary 7.2. For $j = 1, \dots, d$, let X_j be a metric space, $U_j \subset X_j$ open, $C, \alpha > 0$, and let f be a function on $U_1 \times \dots \times U_d$ such that for any $j = 1, \dots, d$ and any $x_i \in U_i$ with $i \neq j$, the function

$$y \mapsto f(x_1, \dots, x_{j-1}, y, x_{j+1}, \dots, x_d)$$

is (C, α) -good on U_j . Then f is $(dC, \alpha/d)$ -good on $U_1 \times \dots \times U_d$.

In particular, polynomials of degree $\leq n$ in d variables are $(dC_n, 1/dn)$ -good on \mathbb{R}^d , where C_n is the constant computed in the previous section.

Is it optimal? I doubt it very much, for example because of the next problem:

Problem 7.6. Show that any polynomial of degree ≤ 1 in d variables (that is, any linear function) is $(\tilde{C}_d, 1/d)$ -good on \mathbb{R}^d for some constant \tilde{C}_d dependent only on d . In other words when the degree of polynomials is 1, the exponent does not go down with extra variables!

Problem 7.7. What happens for quadratic polynomials in 2 variables? Can you improve the estimate with $\alpha = 1/4$?

8. MORE ON (C, α) -GOOD FUNCTIONS

Homework problems due: Monday, December 12

(These two sections contain two parts of the last homework)

A follow-up on Problem 6.6: we would like to show in this section that if we replace a polynomial of degree n (characterized by the property that its n th derivative is constant) with a function whose n th derivative is almost constant, that is, admits uniform bounds from both

sides, it will still have the (C, α) -good property, with the constants depending only on the bounds for the derivative. As in the previous section, the results will be far from optimal and the reader is invited to try improving them, in a series of problems.

We start with a perturbation of quadratic polynomials.

Theorem 8.1. *Let $V \subset \mathbb{R}$ be open, and let $f \in C^2(V)$ be such that for some constants $0 < a \leq A$ one has*

$$a \leq |f''(x)| \leq A \quad \forall x \in V. \quad (8.1)$$

Then f is $(2\sqrt{22A/a}, 1/2)$ -good on V .

Proof. We present a simplification of the argument from [KM]. Fix a subinterval B of V and denote by b the length of B and by s the supremum of $|f|$ on B . Take $\varepsilon > 0$; since, by the lower estimate in (8.1), the second derivative of f does not vanish on B , the set $B^{f,\varepsilon}$ (recall the notation introduced in the previous section) consists of at most 2 intervals.

Problem 8.1. Prove it.

Let I be the maximal of those, and denote its length by r . Then

$$|B^{f,\varepsilon}| \leq 2r, \quad (8.2)$$

so it suffices to estimate r from above.

Lemma 8.2. $r \leq 2\sqrt{6\varepsilon/a}$.

Proof. Divide I into 2 equal parts by points x_1, x_2, x_3 , and let P be the Lagrange polynomial of degree 2 formed by using values of f at these points, i.e. given by the expression in the right hand side of (6.3) with $n = 2$. Then there exists $x \in I$ such that $P''(x) = f''(x)$.

Problem 8.2. Prove it.

Hence, by the lower estimate in (8.1), $|P''(x)| \geq a$. On the other hand, one can differentiate the right hand side of (6.3) twice to get

$$|P''(x)| \leq 3\varepsilon \frac{2}{(r/2)^2} = 24\varepsilon/r^2.$$

Combining the last two inequalities yields the desired estimate. \square

Recall that, since we are after the (C, α) -good property, we would like to have an upper estimate for r in the form $r \leq C(\varepsilon/s)^\alpha b$. Thus let us rewrite the conclusion of the lemma as

$$r \leq 2\sqrt{\frac{6s}{ab^2}} \left(\frac{\varepsilon}{s}\right)^{1/2} b = 2\sqrt{\frac{6t}{a}} \left(\frac{\varepsilon}{s}\right)^{1/2} b, \quad (8.3)$$

where we introduced a parameter $t \stackrel{\text{def}}{=} s/b^2$. We see that the above estimate is good when t is small, and to finish the proof it suffices to produce an estimate improving (8.3) for large values of t . Here it goes:

Lemma 8.3. $r \leq \sqrt{\frac{10A/a}{1-A/2t}} \cdot \left(\frac{\varepsilon}{s}\right)^{1/2} b$.

Proof. Let Q be the affine approximation to f (the Taylor polynomial of f of degree 1) at x_1 . By Taylor's formula,

$$|f(x_2) - Q(x_2)| \leq \|f''\|_I \frac{(r/2)^2}{2} \stackrel{(8.1)}{\leq} \frac{Ar^2}{8} \stackrel{\text{Lemma 8.2}}{\leq} \frac{A}{8} \frac{24\varepsilon}{a} = 3\frac{A}{a}\varepsilon.$$

But also $|f(x_2)| \leq \varepsilon$, therefore

$$|Q(x_2)| \leq \left(3\frac{A}{a} + 1\right) \varepsilon \stackrel{\text{to simplify computations}}{\leq} 4\frac{A}{a}\varepsilon.$$

We now apply Lagrange's formula to reconstruct Q on B by its values at x_1, x_2 . As in the proof of Proposition 6.2, we get

$$\|Q\|_B \leq \left(4\frac{A}{a}\varepsilon + \varepsilon\right) \frac{b}{r/2} \stackrel{\text{to simplify computations}}{\leq} 10\frac{A}{a} \cdot \varepsilon \frac{b}{r}. \quad (8.4)$$

Finally, the difference between f and Q on B is, again by the upper estimate in (8.1), bounded from above by $Ab^2/2$, so from (8.4) one deduces that

$$s \leq 10\frac{A}{a} \cdot \varepsilon \frac{b}{r} + Ab^2/2 \stackrel{\text{to simplify computations}}{\leq} 10\frac{A}{a} \cdot \varepsilon \frac{b^2}{r^2} + Ab^2/2,$$

or, equivalently,

$$r \leq \sqrt{\frac{10\frac{A}{a} \cdot \varepsilon b^2}{s - Ab^2/2}} = \sqrt{\frac{10\frac{A}{a}}{1 - Ab^2/2s}} \cdot \left(\frac{\varepsilon}{s}\right)^{1/2} b.$$

which is what we wanted to prove. \square

Problem 8.3. Show that the right hand sides of the two inequalities in Lemmas 8.2 and 8.3 are equal to each other when $t = 11A/12$.

Using the result of the above problem, we substitute $t = 11A/12$ in (8.3) to obtain $r \leq 2\sqrt{\frac{11A}{2a}} \left(\frac{\varepsilon}{s}\right)^{1/2} b$, which, in view of (8.2), gives the conclusion of the theorem. \square

Problem 8.4. Repeat the above argument under an assumption that f is n times continuously differentiable function with (8.1) replaced by

$$a \leq |f^{(n)}(x)| \leq A \quad \forall x \in V. \quad (8.5)$$

This is not hard at all, one simply needs to use Lagrange and Taylor polynomials of higher order. Obtain an explicit estimate of C in terms of n and A/a .

Problem 8.5. One can see that as $A/a \rightarrow 1$, the conclusion of Theorem 8.1 does not tend to that of Theorem 6.2. The same of course happens with the answer to the above problem. How can one improve the argument?

- (a) A lame attempt would be not being generous in several places in the above proof labelled “to simplify computations”. Most likely “not simplifying computations” won’t produce any reasonable improvement, but feel free to try and see what happens to be a more precise answer.
- (b) Try to think about what happens if f in Theorem 8.1 was assumed to be a quadratic polynomial to begin with. Where exactly does the loss occur? Is there anything one can do to fight with it? Maybe it is possible to prove a Lagrange-type estimate (as opposed to the exact equality of the Lagrange interpolation formula) for functions satisfying (8.1) or (8.5)?

REFERENCES

- [DM] S. G. Dani and G. A. Margulis, *Limit distributions of orbits of unipotent flows and values of quadratic forms*, Adv. in Soviet Math. **16** (1993), 91–137.
- [E] A. Engel, **Problem-Solving Strategies**, Springer-Verlag, New York, 1998.
- [KM] D. Kleinbock and G. A. Margulis, *Flows on homogeneous spaces and Diophantine approximation on manifolds*, Ann. Math. **148** (1998), 339–360.
- [KT] D. Kleinbock and G. Tomanov, *Flows on S -arithmetic homogeneous spaces and applications to metric Diophantine approximation*, Preprint (2004).

DEPARTMENT OF MATHEMATICS, BRANDEIS UNIVERSITY, WALTHAM, MA 02454-9110

E-mail address: `kleinboc@brandeis.edu`

URL: `http://people.brandeis.edu/~kleinboc/47a/intro.html`