KTH Computer Science and Communication

Algebraic Gems in TCS: Problem Set 3

Due: Friday Jan 16, 2015, at 23:59. Submit your solutions as a PDF file by e-mail to `jakobn at kth dot se` with the subject line `Problem set 3: ⟨your full name⟩`. Name the PDF file `PS3_⟨YourFullName⟩.pdf` (with your name coded in ASCII without national characters), and also state your name and e-mail address at the top of the first page. Solutions should be written in LaTeX or some other math-aware typesetting system. Please try to be precise and to the point in your solutions and refrain from vague statements. *Write so that a fellow student of yours can read, understand, and verify your solutions.* In addition to what is stated below, the general rules stated on the course webpage always apply.

Collaboration: Discussions of ideas in groups of two people are allowed—and indeed, encouraged—but you should write down your own solution individually and understand all aspects of it fully. You should also acknowledge any collaboration. State at the beginning of the problem set if you have been collaborating with someone and if so with whom. (Note that collaboration is on a per problem set basis, so you should not discuss different problems on the same problem set with different people.)

Reference material: Some of the problems are "classic" and hence it might be easy to find solutions on the Internet, in textbooks or in research papers. It is not allowed to use such material in any way unless explicitly stated otherwise. Anything said during the lectures or in the lecture notes should be fair game, though, unless you are specifically asked to show something that we claimed without proof in class. It is hard to pin down 100% formal rules on what all this means—when in doubt, ask the lecturer.

About the problems: Some of the problems are meant to be quite challenging and you are not necessarily expected to solve all of them. A total score of around 100 points should be enough for grade E, 150 points for grade D, 200 points for grade C, 250 points for grade B, and 300 points for grade A on this problem set. Any corrections or clarifications will be given at `piazza.com/kth.se/fall2014/dd2442/` and any revised versions will be posted on the course webpage `www.csc.kth.se/DD2442/semteo14/`.

1 (20 p) For non-negative integer vectors $\mathbf{i} = (i_1, \ldots, i_n)$ and $\mathbf{j} = (j_1, \ldots, j_n)$ and variable vectors $\mathbf{x} = (x_1, \ldots, x_n)$ and $\mathbf{y} = (y_1, \ldots, y_n)$, let $\mathbf{x}^\mathbf{i} = \prod_{\ell=1}^n x_\ell^{i_\ell}$, $\binom{\mathbf{i}}{\mathbf{j}} = \prod_{\ell=1}^n \binom{i_\ell}{j_\ell}$, and $\mathbf{x} + \mathbf{y} = (x_1 + y_1, \ldots, x_n + y_n)$. Show that the coefficient of $\mathbf{x}^\mathbf{j}\mathbf{y}^{\mathbf{i}-\mathbf{j}}$ in $(\mathbf{x}+\mathbf{y})^\mathbf{i}$ equals $\binom{\mathbf{i}}{\mathbf{j}}$.
Remark: Note that we want the solution not only to claim that this equality holds (this is already in the problem statement) but to *explain why* or at least *explicitly show that* this is indeed true, and to do so in a convincing manner.

2 (40 p) Let $\mathbf{x} = (x_1, \ldots, x_n)$ and $\mathbf{z} = (z_1, \ldots, z_n)$ be vectors of variables and $\mathbf{i} = (i_1, \ldots, i_n)$ a vector of non-negative integers, and let $wt(\mathbf{i}) = \sum_{\ell=1}^n i_\ell$. Define the $\mathbf{i}$*th Hasse derivative* of $p(\mathbf{x}) \in \mathbb{F}[\mathbf{x}]$, denoted $p^{(\mathbf{i})}(\mathbf{x})$, to be the coefficient of $\mathbf{z}^\mathbf{i}$ in the polynomial $p(\mathbf{x}+\mathbf{z}) \in \mathbb{F}[\mathbf{x}, \mathbf{z}]$. Prove the following basic properties of Hasse derivatives:

2a For polynomials $p(\mathbf{x}), q(\mathbf{x}) \in \mathbb{F}[\mathbf{x}]$ it holds that $p^{(\mathbf{i})}(\mathbf{x}) + q^{(\mathbf{i})}(\mathbf{x}) = (p+q)^{(\mathbf{i})}(\mathbf{x})$.

2b If $p(\mathbf{x}) \in \mathbb{F}[\mathbf{x}]$ is homogeneous of degree d, then either $p^{(\mathbf{i})}(\mathbf{x})$ is homogeneous of degree $d - wt(\mathbf{i})$ or $p^{(\mathbf{i})}(\mathbf{x}) = 0$.

2c Let $H_p(\mathbf{x})$ denote the homogeneous part of $p(\mathbf{x}) \in \mathbb{F}[\mathbf{x}]$ of highest total degree. Then either $\big(H_p\big)^{(\mathbf{i})}(\mathbf{x}) = \big(H_{p^{(\mathbf{i})}}\big)(\mathbf{x})$ or $\big(H_p\big)^{(\mathbf{i})}(\mathbf{x}) = 0$.

2d For $p(\mathbf{x}) \in \mathbb{F}[\mathbf{x}]$ it holds that $\big(p^{(\mathbf{i})}\big)^{(\mathbf{j})}(\mathbf{x}) = \binom{\mathbf{i}+\mathbf{j}}{\mathbf{i}} p^{(\mathbf{i}+\mathbf{j})}(\mathbf{x})$.

3 (30 p) Looking back at our lectures on Kakeya sets in finite fields, we spent quite some time on defining the Hasse derivative and exploring its properties, but once we got started on the proof of the (almost) optimal Kakeya set size bound it seems we somehow never really used that we had Hasse derivatives instead of standard derivatives (except that it made our lives significantly more complicated). This raises the obvious question of whether we actually needed to go through all the pain of Hasse derivatives in the first place. Clearly, there must be some reason they are there in [DKSS13], but this paper also contains a number of other results on Kakeya-ish sets for curves and statistical Kakeya sets and what have you, and so maybe that was where the Hasse derivative was really needed.

So... The moment of truth: Did we ever actually need the Hasse derivatives for the proofs to go through? Or would the standard derivative have worked equally well for the limited sets of results in the beginning of [DKSS13] that we covered in class? Answer these questions, and back up your answer with a formal argument.

4 (40 p) For a function $f : \mathbb{F}_2^n \to \mathbb{F}_2$ and vectors $\mathbf{y}_1, \ldots, \mathbf{y}_\ell \in \mathbb{F}_2^n$, let

$$T_f(\mathbf{y}_1, \ldots, \mathbf{y}_\ell) = \sum_{\emptyset \neq I \subseteq [\ell]} f\big(\textstyle\sum_{i \in I} \mathbf{y}_i\big) \ . \tag{1}$$

Prove that f is a degree-d polynomial with constant term 0 if and only if for all $\mathbf{y}_1, \ldots, \mathbf{y}_{d+1} \in \mathbb{F}_2^n$ it holds that $T_f(\mathbf{y}_1, \ldots, \mathbf{y}_{d+1}) = 0$.

5 (20 p) Let $f : \mathbb{F}_2^n \to \mathbb{F}_2$ be a function and $\mathbf{y} \in \mathbb{F}_2^n$ a fixed vector. For uniformly and independently sampled random vectors $\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_{d+1} \in \mathbb{F}_2^n$, define

$$T_f^{\mathbf{y}}(\mathbf{y}_2, \ldots, \mathbf{y}_{d+1}) = T_f(\mathbf{y}, \mathbf{y}_2, \ldots, \mathbf{y}_{d+1}) + f(\mathbf{y}) \ , \tag{2}$$

with $T_f(\mathbf{y}, \mathbf{y}_2, \ldots, \mathbf{y}d + 1)$ given by (1);

$$g(\mathbf{y}) = \begin{cases} 1 & \text{if } \Pr_{\mathbf{y}_2, \ldots, \mathbf{y}_{d+1}}\big[T_f^{\mathbf{y}}(\mathbf{y}_2, \ldots, \mathbf{y}_{d+1}) = 1\big] \geq 1/2, \\ 0 & \text{otherwise,} \end{cases} \tag{3}$$

and

$$\eta = \Pr_{\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_{d+1}}\big[T_f(\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_{d+1}) \neq 0\big] \ , \tag{4}$$

where we recall that η in (4) is the probability of detecting that f is not a degree-d polynomial and that g in (3) is the "majority-vote" function trying to correct f to pass the polynomiality test. Show that the distance between f and g is $\delta(f, g) \leq 2\eta$.

6 (40 p) In Lectures 18–19 we proved the [AKKLR05] result for testing of degree-d polynomials with constant term 0. Then, when starting to discuss the paper [BKSSZ09] in Lecture 20, we happily claimed that in fact the same analysis would sort of work for any degree-d polynomials regardless of what the constant term is. Your task is now to verify this claim.

To this end, go over our exposition of the result in [AKKLR05] carefully, and for every definition, lemma, theorem, or similar explain if and how the statement needs to be modified, what if any modifications are needed in the proofs, and how the parameters of the low-degree tester are affected as a result of this.

Remark: You do *not* need to provide a fully written out proof of [AKKLR05] for general degree-d polynomials, but you should provide enough details so that a fellow students of yours could in principle produce such a detailed write-up based on the exposition we did in class and your description of the required modifications. In particular, if a lemma goes through without any modification whatsoever, then just stating so is fine, provided that a short explanation is also given why no change is needed. Analogously, if some proof needs to be adjusted in some way, then a brief explanation of how the adjustment should be made is sufficient—there is no need to repeat an entire proof.

7 (30 p) Another discrepancy between [AKKLR05] and [BKSSZ09] regarding the description of the test was whether it was guaranteed to pick a subspace of maximal dimension or not. Again, we handwaved that this did not really make much of difference. Formally, let $f : \mathbb{F}_2^n \to \mathbb{F}_2$ be a function and consider the following two tests:

Test $T_{d,\ell}$: Pick uniformly and independently at random a full-rank matrix $M \in \mathbb{F}_2^{n\times\ell}$ and a vector $\mathbf{b} \in \mathbb{F}_2^n$, let $A = \{M\mathbf{x} + \mathbf{b} \big| x \in \mathbb{F}_2^\ell\}$, and accept f if and only if $f\restriction_A$ is a degree-d polynomial.

Test $T_{d,\leq\ell}$: Same as above, except $M \in \mathbb{F}_2^{n\times\ell}$ is picked uniformly and independently at random among all matrices regardless of rank.

Prove that if $n \geq d + 1$, then

$$\Pr\big[T_{d,d+1} \text{ rejects } f\big] \;\geq\; \Pr\big[T_{d,\leq d+1} \text{ rejects } f\big] \;\geq\; \frac{1}{4}\Pr\big[T_{d,d+1} \text{ rejects } f\big] \;.$$

8 (30 p) Recall that a hyperplane in $\mathbb{F}_2^n$ is a set of points $A = \{\mathbf{x} \in F_2^n \mid L_\mathbf{c}(\mathbf{x}) = b\}$ for some $b \in \mathbb{F}_2$ and some $L_\mathbf{c}(\mathbf{x}) = \langle \mathbf{c}, \mathbf{x}\rangle = \sum_{i=1}^n c_i x_i$ for $\mathbf{c} \in \mathbb{F}_2$. We say that the hyperplanes $A_1, \ldots, A_\ell$ are linearly independent if the corresponding linear parts $L_{\mathbf{c}_1}, \ldots, L_{\mathbf{c}_\ell}$ are linearly independent. Prove the following basic properties of hyperplanes:

8a There are exactly $2^{n+1} - 2$ distinct hyperplanes in $\mathbb{F}_2^n$.

8b Among any $2^\ell - 1$ distinct hyperplanes there are at least ℓ linearly independent hyperplanes.

8c For any linearly independent hyperplanes $A_1, \ldots, A_\ell$ there is an affine, invertible transform that sends A_i to the hyperplane $\{\mathbf{y} \mid y_i = 0\}$ for $i = 1, \ldots, n$ (i.e., an invertible matrix $M \in \mathbb{F}_2^{n\times n}$ and a $\mathbf{b} \in \mathbb{F}_2^n$ such that $\mathbf{x}$ is sent to $\mathbf{y} = M\mathbf{x} + \mathbf{b}$).

9 (150 p) In our very final lecture, we omitted proofs for a number of important claims in the exposition of [BKSSZ09]. Your task in this problem is to fill in the missing details to establish these claims.

9a Prove for the test $T_{d,\ell}$ in Problem 7 that if $k \geq \ell$, then $\Pr\left[T_{d,k} \text{ rejects } f\right] \geq \Pr\left[T_{d,\ell} \text{ rejects } f\right]$.

9b Pick uniformly and independently at random a full-rank matrix $M \in \mathbb{F}_2^{n \times \ell}$ and a vector $\mathbf{b} \in \mathbb{F}_2^n$, and denote $\mathbf{a_x} = M\mathbf{x} + \mathbf{b}$ for $\mathbf{x} \in \mathbb{F}_2^\ell$. Prove that for any fixed $\mathbf{x}, \mathbf{y} \in \mathbb{F}_2^\ell$, $\mathbf{x} \neq \mathbf{y}$, it holds that $\mathbf{a_x}$ is uniformly distributed over $\mathbb{F}_2^n$ and $\mathbf{a_y}$ is uniformly distributed over $\mathbb{F}_2^n \setminus \{\mathbf{a_x}\}$.

9c Let $p(\mathbf{x}) \in \mathbb{F}_2[\mathbf{x}]$ be a polynomial and let $\mathbf{y} = M\mathbf{x} + \mathbf{b}$ be an invertible affine transform on $\mathbb{F}_2^n$. Prove that the polynomial $q(\mathbf{x}) \in \mathbb{F}_2[\mathbf{x}]$ defined by $q(\mathbf{x}) = p(M\mathbf{x} + \mathbf{b})$ has exactly the same total degree as $p(\mathbf{x})$. When and why is this important in the proofs in [BKSSZ09]?

9d Recall that a *k-flat* is an affine subspace of $\mathbb{F}_2^n$ of dimension k. Consider the following two experiments:

1. Pick a uniformly random k-flat $B \subseteq \mathbb{F}_2^n$.
2. Pick a uniformly random hyperplane $C' \subseteq \mathbb{F}_2^n$ and then a uniformly random k-flat $C \subseteq C'$.

Show that the k-flats B and C have exactly the same distribution.

9e Show that any function $f : \mathbb{F}_2^n \to \mathbb{F}_2$ has a *unique* representation as a multilinear polynomial in $\mathbb{F}_2[\mathbf{x}]$ (i.e., , a polynomials with all individual variable degrees being 0 or 1).

9f Recall that two hyperplanes A and A' in $\mathbb{F}_2^n$ are *complementary* if $A \cup A' = \mathbb{F}_2^n$. Prove that if $A_1, \ldots, A_\ell$ are linearly independent hyperplanes for $\ell > d$ and A is a hyperplane that is not complementary to any A_i, then it holds that

$$\frac{\left|A \cap \bigcup_{j=1}^{\ell} A_i\right|}{|A|} > 1 - 2^{-d} \ .$$

Actually, looking more closely at this part of the proof in [BKSSZ09], it seems we do *not* have the guarantee that A is not complementary to any A_i. What happens to the above claim in that case? Is it still true that Sublemma 7 in our notes (Claim 15 in [BKSSZ09]) holds? Or if it does not, is there some way to fix the argument? Or do we have a guarantee that A is not complementary to any A_i after all?

Remark: Solving this subproblem completely yields about a third of the total amount of points on this problem.

9g For a fixed hyperplane $A_i \subseteq \mathbb{F}_2^n$ and a uniformly random $\mathbf{z} \in \mathbb{F}_2^n$, define the random variable

$$Y_i = Y_i(\mathbf{z}) = \begin{cases} 1 & \text{if } \mathbf{z} \in A_i, \\ 0 & \text{otherwise.} \end{cases}$$

Prove that if $A_i \neq A_j$ are not complementary hyperplanes, then Y_i and Y_j are independent.