An Entropic Proof of Chang's Inequality

Russell Impagliazzo* Cristopher Moore† Alexander Russell‡

May 17, 2012

Abstract

Chang's lemma is a useful tool in additive combinatorics and the analysis of Boolean functions. Here we give an elementary proof using entropy. The constant we obtain is tight, and we give a slight improvement in the case where the variables are highly biased.

1 The lemma

For $S \in \{0,1\}^n$, let $\chi_k : \{\pm 1\}^n \to \mathbb{R}$ denote the character

$$\chi_S(x) = \prod_{i \in S} x_i \,.$$

For any function $f : \{\pm 1\}^n \to \mathbb{R}$, we can then define its Fourier transform $\widehat{f} : \{0,1\}^n \to \mathbb{R}$ as

$$\widehat{f}(S) = \mathop{\mathbb{E}}_x f(x)\chi_S(x) = \frac{1}{2^n} \sum_x f(x)\chi_S(x) \,.$$

For characters of Hamming weight 1, we will abuse notation by writing $\widehat{f}(i)$ instead of $\widehat{f}(\{i\})$.

Chang's lemma [1, 2] places an upper bound on the total Fourier weight, i.e., the sum of $\widehat{f}^2$, of the characteristic function of a small set on the characters with Hamming weight one.

Lemma 1. *Let $A \subseteq \{\pm 1\}^n$ such that $|A| = 2^n \alpha$, and let $f = \mathbb{1}_A$ be its characteristic function. Then*

$$\sum_{i=1}^n \widehat{f}(i)^2 \le 2\alpha^2 \ln \frac{1}{\alpha} \,.$$

Proof. Suppose that we sample x according to the uniform distribution on A. Since the mutual information is nonnegative, the entropy $H(x)$ is at most the sum of the entropies of the individual bits,

$$H(x) \le \sum_{i=1}^n H(x_i) \,.$$

*russell@cs.ucsd.edu, Department of Computer Science and Engineering, University of California San Diego
†moore@santafe.edu, Department of Computer Science, University of New Mexico and Santa Fe Institute
‡acr@cse.uconn.edu, Department of Computer Science and Engineering, University of Connecticut

This gives

$$n \ln 2 + \ln \alpha \le \sum_{i=1}^{n} h(p_i^+) \tag{1}$$

where p_i^+ denotes the probability that $x_i = +1$,

$$p_i^+ = \frac{1}{2}\left(1 + \mathop{\mathbb{E}}_{x \in A} x_i\right) = \frac{1}{2}\left(1 + \frac{\widehat{f}(i)}{\alpha}\right).$$

and where h denotes the entropy function

$$h(p) = -p \ln p - (1-p)\ln(1-p)\,.$$

The Taylor series around $p = 1/2$ gives

$$h\left(\frac{1+x}{2}\right) = \ln 2 \; - \sum_{t=2,4,6,\dots} \frac{x^t}{t(t-1)} \le \ln 2 - \frac{x^2}{2}\,, \tag{2}$$

so (1) becomes

$$\ln \alpha \le -\frac{1}{2} \sum_{i=1}^{n} \frac{\widehat{f}(i)^2}{\alpha^2}\,,$$

Rearranging completes the proof. □

2 Variations

The lemma (and our proof) apply equally well to the Fourier weight $\sum_{S \in B} \widehat{f}(S)^2$ of any basis B of $\mathbb{F}_2^n$, since the set of parities $\{\prod_{i \in S} x_i \mid S \in B\}$ determines x. This gives the following commonly-quoted form of Chang's lemma.

Lemma 2. *Let $A \subseteq \{\pm 1\}^n$ such that $|A| = 2^n \alpha$, and let $f = \mathbb{1}_A$ be its characteristic function. Fix $\rho > 0$ and let $R \subset \mathbb{F}_2^n$ be the set $\{S : |\widehat{f}(S)| > \rho\alpha\}$. Then R spans a space of dimension less than $d = 2\rho^{-2} \ln(1/\alpha)$.*

Proof. If R spans a space of dimension d or greater, there is a set of d linearly independent vectors in R. Completing to form a basis B gives $\sum_{S \in B} \widehat{f}(S)^2 > 2\alpha^2 \ln(1/\alpha)$, violating Lemma 1. □

For any integer $k \ge 1$, there are bases consisting entirely of vectors of Hamming weight k. Fixing k and averaging over all such bases gives

$$\sum_{S:|S|=k} \widehat{f}(S)^2 \le \frac{2}{n}\binom{n}{k} \alpha^2 \ln \frac{1}{\alpha} \le \frac{2n^{k-1}}{k!}\, \alpha^2 \log(1/\alpha)\big)\,.$$

This also follows immediately from Shearer's lemma. However, this is noticeably weaker than the "weight k bound"

$$\sum_{S:|S|=k} \widehat{f}(S)^2 = O\big(\alpha^2 \log^k(1/\alpha)\big)\,.$$

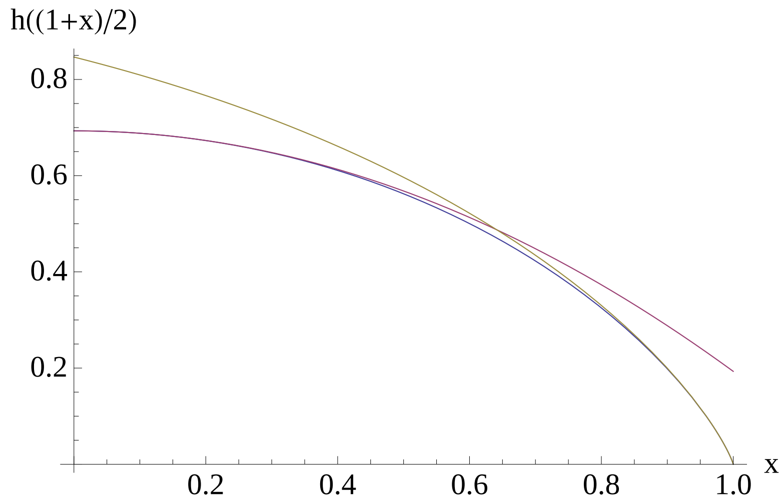

Figure 1: The entropy function $h(p)$ where $p = (1+x)/2$ and $x \leq 0 \leq 1$, with the upper bounds (2) (which is tight when $|x|$ is small) and (3) (which is tight when $|x|$ is close to 1).

Finally, we note that if some bits are highly biased, i.e., if $|\widehat{f}(i)|/\alpha$ is close to 1, we can replace (2) with the bound

$$h(p) \leq p(1 - \ln p)\,, \tag{3}$$

which is tight when p is small. Combining this with the corresponding bound for p close to 1 gives

$$h\left(\frac{1+x}{2}\right) \leq \frac{1-|x|}{2}\left(1 - \ln\frac{1-|x|}{2}\right)\,.$$

We compare this bound with (2) in Figure 1. This gives another version of Lemma 1:

Lemma 3. *Let $A \subseteq \{\pm 1\}^n$, let $f = \mathbb{1}_A$ be its characteristic function, and let*

$$\delta_i = \frac{1}{2}\left(1 - \frac{|\widehat{f}(i)|}{\alpha}\right) = \min\left(p_i^+, 1 - p_i^+\right)\,.$$

Then

$$\sum_{i=1}^{n} \delta_i\left(1 - \ln \delta_i\right) \geq \ln |A|\,. \tag{4}$$

This is nearly tight, for instance, if A is the set of vectors with Hamming weight 1. Then $|A| = n$, $\delta_i = 1/n$, and (4) reads $1 + \ln n \geq \ln n$.

Acknowledgments

We thank Ryan O'Donnell for a wonderful set of lectures on the analysis of Boolean functions at the Bellairs Research Institute, and Ran Raz for helpful communications. C.M. and A.R. are supported by NSF grant CCF-1117426 and ARO contract W911NF-04-R-0009.

References

[1] Michel Talagrand, “How much are increasing sets positively correlated?” *Combinatorica* 16 (2) 243–258, 1996.

[2] Mei-Chu Chang, “A polynomial bound in Freiman’s theorem.” *Duke Math. J.* 113(3) 399–419, 2002.