ISRAEL JOURNAL OF MATHEMATICS **131** (2002), 269–276

ON THE DISTRIBUTION OF THE FOURIER SPECTRUM OF BOOLEAN FUNCTIONS

BY

J. BOURGAIN

School of Mathematics, Institute for Advanced Study
Princeton, NJ 08540, USA
e-mail: bourgain@math.ias.edu

ABSTRACT

In this paper we obtain a general lower bound for the tail distribution of the Fourier spectrum of Boolean functions f on $\{1,-1\}^N$. Roughly speaking, fixing $k \in \mathbb{Z}_+$ and assuming that f is not essentially determined by a bounded number (depending on k) of variables, we have that $\sum_{|S|>k} |\hat{f}(S)|^2 \gtrsim k^{-1/2-\varepsilon}$. The example of the majority function shows that this result is basically optimal.

Introduction

Over recent years, a new area in Harmonic Analysis has emerged, which is the Fourier Analysis of Boolean functions $f\colon \{1,-1\}^I \to \{0,1\}$.

Motivated by problems from complexity theory and computer science, a number of remarkable results were obtained from the study of the Fourier transform $\hat{f}$ of f. In this context we mention, for instance, the works of Kahn, Kalai and Linial [KKL] on the influence of variables and E. Friedgut [Fr] on the characterization of the sharp threshold of monotone properties. They rely crucially on the analysis of the Fourier transform.

There is a general philosophy which claims that if f defines a property of 'high complexity', then supp $\hat{f}$, the support of the Fourier transform, has to be 'spread out'. The result in this paper is one more illustration of this phenomenon: If f

Received January 31, 2001

is not essentially determined by a few variables, then the tail distribution of $\hat{f}$ satisfies a lower bound

$$\sum_{|S|>k} |\hat{f}(S)|^2 \gg c_\varepsilon k^{-\frac{1}{2}-\varepsilon}$$

for all (fixed) k. A precise formulation appears below. This estimate, which turns out to be basically sharp, thus expresses to what extent $\hat{f}$ may be fully concentrated on coefficients $\hat{f}(S)$ of low weight $|S|$. For a real function f on $\{-1,1\}^N$ let $f = \sum \hat{f}(S) w_S$ be its Fourier expansions. Here,

$$w_S(x_1, x_2, \dots, x_N) = (-1)^{\sum_{I\in S} x_i}.$$

L^2-weight of the tail of the Fourier spectrum

The main result of this Note is the following

PROPOSITION*: *Let* $f = \chi_A$, $A \subset \{1,-1\}^N$. *Let* $k > 0$ *be an integer and* $\gamma > 0$ *a fixed constant. Assume*

$$\sum \{ |\hat{f}(S)|^2 \Big| |\hat{f}(S)| < \gamma 4^{-k^2} \} > \gamma^2. \tag{1}$$

Then

$$\sum_{|S|>k} |\hat{f}(S)|^2 \gtrsim k^{-\frac{1}{2}-}.^{**} \tag{2}$$

Proof: We may clearly assume that

$$\sum_{|S|>k} |\hat{f}(S)|^2 < \frac{1}{100}\gamma^2. \tag{3}$$

Fix $0 < \kappa < 1$ and define

$$I_0 = \Big\{ i \in [1,N] \Big| \sum_{i\in S, |S|\le k} |\hat{f}(S)|^2 > \kappa \Big\}.$$

Then

$$\kappa |I_0| \le \sum_{i=1}^{N} \sum_{i\in S, |S|\le k} |\hat{f}(S)|^2 < k \quad \text{and} \quad |I_0| < \kappa^{-1} k.$$

* This question was raised by J. Håstad, who obtained a lower estimate of the form C^{-k} in (2). The author is also grateful to G. Kalai for several discussions on this and related topics.

** More precisely, there is the lower bound $c_\varepsilon k^{-1/2-\varepsilon}$ for all $\varepsilon > 0$. This abbreviated notation will be used repeatedly in the sequel.

Thus

$$\sum\left\{|\hat{f}(S)|^2\Big|S\subset I_0,|S|\leq k,|\hat{f}(S)|<\gamma 4^{-k^2}\right\}<(\kappa^{-1}k)^k\gamma^2 16^{-k^2}<\frac{\gamma^2}{100} \tag{4}$$

if we assume

$$(\kappa^{-1}k)^k 16^{-k^2}<1/100. \tag{5}$$

Denote

$$I_0'=[1,N]\backslash I_0.$$

It follows from (1), (3) and (4) that

$$\sum_{\substack{S\cap I_0'\neq\phi\\|S|\leq k}}|\hat{f}(S)|^2>\gamma^2-\frac{1}{100}\gamma^2-\frac{1}{100}\gamma^2>\frac{1}{2}\gamma^2. \tag{6}$$

Define for $t\geq 0$

$$\rho_t=\sum_{2^t\leq|S\cap I_0'|<2^{t+1}}|\hat{f}(S)|^2 \tag{7}$$

so that (6) implies that

$$\sum_{0\leq t\leq\log k}\rho_t>\gamma^2/2 \tag{8}$$

(where $\log k={}^2\log k$).

Next, fix a subset

$$I_1\subset I_0'.$$

Write the variable $x\in\{1,-1\}^N$ as $x=(x_1,x_2)$ with $x_1\in\{1,-1\}^{I_1}$. For a fixed x_2 write $f_{x_2}(x_1)$ for $f(x_2,x_1)$ and write also $F_T(x_2)$ for $\hat{f}_{x_2}(T)$. Thus,

$$f(x_1,x_2)=\sum_{T\subset I_1}F_T(x_2)w_T(x_1).$$

Fix $0<\delta<1$ and $\{\xi_i\}_{i\in I_1}$ independent $\{0,1\}$-valued selectors of mean $1-\delta$. Define

$$I(\omega)=\{i\in I_1|\xi_i(\omega)=1\}.$$

Fix x_2. Since f_{x_2} is $\{0,1\}$-valued,

$$2\int|f_{x_2}-\mathbb{E}_{I(\omega)}[f_{x_2}]|^2dx_1=\int|f_{x_2}-\mathbb{E}_{I(\omega)}[f_{x_2}]|dx_1 \tag{9}$$

where

$$f_{x_2} - \mathbb{E}_{I(\omega)}[f_{x_2}] = \sum_{\substack{T\subset I_1 \\ T\not\subset I(\omega)}} F_T(x_2) w_T(\cdot). \tag{10}$$

Fix $1 < p < 2$. Then by (9) and (10)

$$\begin{aligned}
&(p-1)^{1/2}\Big(\sum_{i\in I_1\setminus I(\omega)} |F_{\{i\}}(x_2)|^2\Big)^{1/2} \\
&\le \|f_{x_2} - \mathbb{E}_{I_{(\omega)}}[f_{x_2}]\|_p \\
&\le \|f_{x_2} - \mathbb{E}_{I(\omega)}[f_{x_2}]\|_1^{1-2/p'} \|f_{x_2} - \mathbb{E}_{I(\omega)}[f_{x_2}]\|_2^{2/p'} \\
&\le \|f_{x_2} - \mathbb{E}_{I(\omega)}[f_{x_2}]\|_2^{2/p} \\
&= \Big[\sum_{\substack{T\subset I_1 \\ T\not\subset I(\omega)}} |F_T(x_2)|^2\Big]^{1/p}
\end{aligned}$$

and

$$\Big[\sum_{i\in I_1}[1-\xi_i(\omega)]|F_{\{i\}}(x_2)|^2\Big]^{p/2} \lesssim (p-1)^{-p/2} \sum_{T\subset I_1} [1-\prod_{i\in T}\xi_i(\omega)]\ |F_T(x_2)|^2. \tag{11}$$

Recall that $\int \xi_i d\omega = 1-\delta$. The left side of (11) is at least

$$\begin{aligned}
&\delta^{p/2}\Big[\sum_{i\in I_1}|F_{\{i\}}(x_2)|^2\Big]^{p/2} - \Big|\sum_{i\in I_1}\big(\xi_i(\omega)-(1-\delta)\big)|F_{\{i\}}(x_2)|^2\Big|^{p/2} \\
\ge&\delta^{p/2}\Big[\sum_{i\in I_1}|F_{\{i\}}(x_2)|^2\Big] - \Big|\sum_{i\in I_1}\big(\xi_i(\omega)-(1-\delta)\big)|F_{\{i\}}(x_2)|^2\Big|^{p/2}.
\end{aligned} \tag{12}$$

Thus

$$\begin{aligned}
\delta^{p/2}\sum_{i\in I_1}|F_{\{i\}}(x_2)|^2 \lesssim&(p-1)^{-p/2}\Big[\sum_{T\subset I_1}[1-\prod_{i\in T}\xi_i(\omega)]|F_T(x_2)|^2\Big] \\
&+\Big|\sum_{i\in I_1}\big(\xi_i(\omega)-(1-\delta)\big)|F_{\{i\}}(x_2)|^2\Big|^{p/2}.
\end{aligned} \tag{13}$$

Integrate (13) in x_2 and ω. Clearly

$$\begin{aligned}
&\int\int\Big|\sum_{i\in I_1}\big(\xi_i(\omega)-(1-\delta)\big)|F_{\{i\}}(x_2)|^2\Big|^{p/2}d\omega dx_2\\
(14)\qquad &\leq\Big[\int\int\Big|\sum_{i\in I_1}\big(\xi_i(\omega)-(1-\delta)\big)|F_{\{i\}}(x_2)|^2\Big|d\omega dx_2\Big]^{p/2}\\
&\lesssim\Big[\int\int\Big[\sum_{i\in I_1}\big(1-\xi_i(\omega)\big)|F_{\{i\}}(x_2)|^4\Big]^{1/2}d\omega dx_2\Big]^{p/2}.
\end{aligned}$$

Estimate

$$\begin{aligned}
|F_{\{i\}}(x_2)|&=\Big|\sum_{S\cap I_1=\{i\}}\hat f(S)w_S(x)\Big|\\
&\leq\Big|\sum_{|S|\leq k,S\cap I_1=\{i\}}\hat f(S)w_S\Big|+\Big|\sum_{|S|>k,S\cap I_1=\{i\}}\hat f(S)w_S\Big|,
\end{aligned}$$

$$\begin{aligned}
\Big(\sum_{i\in I_1}|F_{\{i\}}(x_2)|^4(1-\xi_i(\omega))\Big)^{1/2}\lesssim&\Big(\sum_{i\in I_1}\Big|\sum_{|S|\leq k,S\cap I_1=\{i\}}\hat f(S)w_S\Big|^4\Big)^{1/2}\\
&+\sum_{i\in I_1}\Big|\sum_{|S|>k,S\cap I_1=\{i\}}\hat f(S)w_S\Big|^2(1-\xi_i(\omega)),
\end{aligned}$$

and integrating, by Beckner's inequality

$$\begin{aligned}
&\int\int\Big[\sum_{i\in I_1}|F_{\{i\}}(x_2)|^4(1-\xi_i(\omega))\Big]^{1/2}d\omega dx_2\\
&\leq 3^k\Big[\sum_{i\in I_1}\Big(\sum_{|S|\leq k,S\cap I_1=\{i\}}|\hat f(S)|^2\Big)^2\Big]^{1/2}+\delta\sum_{|S|>k}|\hat f(S)|^2\\
(15)\qquad &\leq 3^k\max_{i\in I_1}\Big(\sum_{\substack{|S|\leq k\\ i\in S}}|\hat f(S)|^2\Big)^{1/2}+\delta\sum_{|S|>k}|\hat f(S)|^2\\
&<3^k\kappa^{1/2}+\delta\sum_{|S|>k}|\hat f(S)|^2
\end{aligned}$$

(we use here the fact that $I_1\cap I_0=\emptyset$ and the definition of I_0).

Substitute (15) in (14). Returning to (13), we thus obtain

$$\delta^{p/2} \sum_{|S\cap I_1|=1} |\hat{f}(S)|^2 \lesssim (p-1)^{-p/2} \sum_S [1-(1-\delta)^{|S\cap I_1|}] |\hat{f}(S)|^2 \tag{16}$$
$$+ \delta^{p/2} \Big(\sum_{|S|>k} |\hat{f}(S)|^2 \Big)^{p/2} + (3^k \kappa^{1/2})^{p/2}.$$

Estimate

$$\begin{aligned} 1-(1-\delta)^{|S\cap I_1|} &\le \delta |S\cap I_1| && \text{if } |S\cap I_0'| \le k,\\ &< 1 && \text{otherwise.} \end{aligned}$$

Thus

$$\delta^{p/2} \sum_{|S\cap I_1|=1} |\hat{f}(S)|^2 \lesssim (p-1)^{-p/2} \delta \sum_{|S\cap I_0'|\le k} |S\cap I_1|\, |\hat{f}(S)|^2 \tag{17}$$
$$+ (p-1)^{-p/2} \sum_{|S|>k} |\hat{f}(S)|^2$$
$$+ \delta^{p/2} \Big(\sum_{|S|>k} |\hat{f}(S)|^2 \Big)^{p/2} + (3^k \kappa^{1/2})^{p/2}.$$

We will now specify the set $I_1 \subset I_0'$.

Fix $0 \le t_0 \le \log k$ and let $I_1 = I_{\omega'}$ be a random subset of I_0' of density $10^{-3} 2^{-t_0}$. This ensures that if $2^{t_0} \le |S\cap I_0'| < 2^{t_0+1}$, then

$$\mathbb{E}_{\omega'}[|S\cap I_1| = 1] > 10^{-4}. \tag{18}$$

Also

$$\mathbb{E}_{\omega'}[|S\cap I_1|] = 10^{-3} 2^{-t_0} |S\cap I_0'|. \tag{19}$$

Applying $\mathbb{E}_{\omega'}$ to (17) and recalling the definition (7) of ρ_t, we get

$$\delta^{p/2} \rho_{t_0} \lesssim (p-1)^{-p/2} \delta \Big(\sum_{t\le \log k} 2^{t-t_0} \rho_t \Big) + (p-1)^{-p/2} \sum_{|S|>k} |\hat{f}(S)|^2 \tag{20}$$
$$+ \delta^{p/2} \Big(\sum_{|S|>k} |\hat{f}(S)|^2 \Big)^{p/2} + (3^k \kappa^{1/2})^{p/2}.$$

In order to have the left term in (20) larger than the first term on the right, take

$$\delta \sim (p-1)^{p/(2-p)} \Big(\frac{2^{t_0} \rho_{t_0}}{\sum_{t\le \log k} 2^t \rho_t} \Big)^{2/(2-p)}. \tag{21}$$

Taking

$$\kappa = 10^{-k} \tag{22}$$

to make the last term in (20) negligible, condition (5) is satisfied and (20) and (21) imply

$$\sum_{|S|>k} |\hat{f}(S)|^2 \gtrsim \min\left\{ (p-1)^{p/(2-p)} \left(\frac{2^{t_0}\rho_{t_0}}{\sum_{t\leq \log k} 2^t \rho_t} \right)^{p/(2-p)} \rho_{t_0}, \rho_{t_0}^{2/p} \right\}. \tag{22}$$

Here $0 \leq t_0 \leq \log k$ and $1 < p < 2$ are arbitrary.

We distinguish two cases.

CASE 1:

$$\sum_{t\leq \log k} 2^t \rho_t < \sqrt{k}. \tag{23}$$

Recalling (8), we may take $0 \leq t_0 \leq \log k$ such that

$$\rho_{t_0} \gtrsim 1/\log k. \tag{24}$$

Take in (22)

$$p = 1 + 1/\log k. \tag{25}$$

From (23), (24) and (25),

$$(22) \gtrsim \min\left((\log k)^{-2} k^{-1/2}, (\log k)^{-2} \right) \gtrsim (\log k)^{-2} k^{-1/2}. \tag{26}$$

CASE 2:

$$\sum_{t\leq \log k} 2^t \rho_t \geq \sqrt{k}.$$

Choose t_0 s.t.

$$2^{t_0}\rho_{t_0} > \frac{1}{\log k} \sum_{t\leq \log k} 2^t \rho_t > \frac{\sqrt{k}}{\log k}, \tag{27}$$

hence

$$\rho_{t_0} > (\log k)^{-1} k^{-1/2}. \tag{28}$$

Take now $p \overset{<}{\longrightarrow} 2$ in (22). We get

$$(22) \gtrsim \min\left((\log k)^{-2/(2-p)-1} k^{-1/2}, (\log k)^{-2} k^{-1/p} \right) > k^{-1/2-}. \tag{29}$$

Thus (2) follows from (26) and (29).

COROLLARY: *Let $f = \mathcal{X}_A, A \subset \{1,1\}^N$ satisfying*

$$|A|(1-|A|) > 1/10. \tag{30}$$

Let $k > 0$ be an integer and assume

$$\max_{|S|\leq k} |\hat{f}(S)| < 4^{-k^2-1}. \tag{31}$$

Then

$$\sum_{|S|>k} |\hat{f}(S)|^2 \gtrsim k^{-1/2-}. \tag{32}$$

Remark: The lower bound (32) in the corollary is basically sharp. This is demonstrated by the example of the 'majority function' which we define as the $\{1,-1\}$-valued function

$$f(\varepsilon) = \text{sign}(\varepsilon_1 + \varepsilon_2 + \cdots + \varepsilon_N) \tag{33}$$

on $\{1,-1\}^N$. It is known (see [K]) that in this case

$$|\hat{f}(S)|^2 \sim \binom{N}{k}^{-1} k^{-3/2} \quad \text{for } |S| = k > 0. \tag{34}$$

Hence

$$\sum_{|S|=k} |\hat{f}(S)|^2 \sim k^{-3/2} \tag{35}$$

and

$$\sum_{|S|>k} |\hat{f}(S)|^2 \sim k^{-1/2}. \tag{36}$$

References

[Fr] E. Friedgut, *Sharp threshold of graph properties, and the k-sat problem*, Journal of the American Mathematical Society **12** (1999), 1017–1054.

[KKL] J. Kahn, G. Kalai and N. Linial, *The influence of variables on Boolean functions*, Proc. 29$^{\text{th}}$ IEEE FOCS 58–80, IEEE, New York, 1988.

[K] M. G. Karpovsky, *Finite Orthogonal Series in the Design of Digital Devices*, Wiley, New York, 1976.