

SIEVE WEIGHTS AND THEIR SMOOTHINGS

ANDREW GRANVILLE, DIMITRIS KOUKOULOPOULOS, AND JAMES MAYNARD

ABSTRACT. We obtain asymptotic formulas for the $2k$ th moments of partially smoothed divisor sums of the Möbius function. When $2k$ is small compared with A , the level of smoothing, the main contribution to the moments come from integers with only large prime factors, as one would hope for in sieve weights. However, if $2k$ exceeds a threshold dependent on A , then the main contribution to the moments come from integers with quite a few prime factors, which is not the intention when designing sieve weights. This threshold occurs when $A = \frac{1}{2k} \binom{2k}{k} - 1$.

One can ask analogous questions for polynomials over finite fields and for permutations, and in these cases the moments behave rather differently, with even less cancellation in the divisor sums. We give, we hope, a plausible explanation for this phenomenon, by studying the analogous sums for Dirichlet characters, and obtaining each type of behaviour depending on whether or not the character is “exceptional”.

CONTENTS

1. Introduction	2
1.1. Some smoothing is necessary to behave like a sieve weight	4
1.2. A heuristic argument	4
1.3. Analogous settings	6
1.4. Two worlds apart and a bridge between them	8
1.5. Further analysis of truncated Möbius divisor sums	11
1.6. Outline of the paper	12
1.7. Notation	13
2. The discrepancy between integers and polynomials	13
2.1. Integer setting	13
2.2. Polynomial analogue	14
2.3. Further analysis	15
2.4. Further obstructions to the analogy	16
3. The analogy for permutations	18
4. The analogy for polynomials over finite fields	21
5. The support of $M_{f_0}(n; R)$	24
6. Inversion formulas	26
7. A combinatorial problem in linear algebra	28
8. Contour integration	31

Date: March 16, 2020.

2010 Mathematics Subject Classification. Primary: 11N35, 11N37; Secondary: 11T06, 20B30, 05A05.

We are grateful to the referee of the paper for an extraordinarily thorough reading of the paper that improved significantly the level of exposition. Our thanks also go to Maksym Radziwill for supplying several key references, and to him, Ben Green and Henryk Iwaniec for helpful discussions. AG is funded by the National Science and Engineering Research Council of Canada and by the European Research Council. DK is funded by the National Science and Engineering Research Council of Canada and by the Fonds de recherche du Québec – Nature et technologies. JM is funded by a Clay research fellowship and a fellowship by examination of Magdalen College, Oxford.

8.1. Initial preparations	32
8.2. Contour shifting	34
8.3. Proof of the auxiliary Lemmas 8.4 and 8.5	37
8.4. Dyadic intervals	43
9. Lower bounds	44
9.1. First manipulations	44
9.2. Mellin transformation	48
9.3. An auxiliary result	49
9.4. Lower bound for the main term	51
9.5. Upper bound for the error term	52
10. On the anatomy of integers contributing to $\mathcal{M}_{f,2k}(R)$	55
10.1. An estimate for the logarithmic means	55
10.2. From logarithmic weights to uniform weights	60
10.3. Estimates for general weight functions	65
11. The analogy for non-exceptional Dirichlet characters	70
12. The analogy for exceptional Dirichlet characters	73
12.1. Initial preparations	73
12.2. The case $k = 1$	74
12.3. Contour shifting	74
12.4. Proof of the auxiliary Lemmas 12.1 and 12.2	76
12.5. Lower bounds	81
References	83

1. INTRODUCTION

Sieve methods are a set of techniques which give upper and lower bounds for the number of elements of a set of integers \mathcal{A} which have no ‘small’ prime factors. Their key benefit is that they are very flexible - one can obtain bounds of the correct order of magnitude for many interesting sets \mathcal{A} , even though obtaining asymptotic formulae looks completely hopeless. In particular, they are typically very effective at obtaining upper bounds for the number of primes in sets \mathcal{A} of interest which are only worse than the conjectured truth by a constant factor.

One of the most important sieves in the Selberg sieve. Selberg’s approach [21] starts with the inequality

$$(1.1) \quad \left(\sum_{\substack{d|n \\ P^+(d) \leq z}} \lambda_d \right)^2 \geq \sum_{\substack{d|n \\ P^+(d) \leq z}} \mu(d) = \begin{cases} 1, & P^-(n) \geq z, \\ 0, & \text{otherwise,} \end{cases}$$

which is valid for any real numbers λ_d with $\lambda_1 = 1$. Here $P^+(n)$ and $P^-(n)$ are the largest and smallest prime factors of n respectively. Summing (1.1) over $n \in \mathcal{A}$ gives

$$\#\{n \in \mathcal{A} : P^-(n) \geq z\} \leq \sum_{n \in \mathcal{A}} \left(\sum_{\substack{d|n \\ P^+(d) \leq z}} \lambda_d \right)^2 = \sum_{P^+(d_1), P^+(d_2) \leq z} \lambda_{d_1} \lambda_{d_2} \cdot \#\{n \in \mathcal{A} : [d_1, d_2] | n\},$$

which is a quadratic form in the variables λ_d . Provided d_1 and d_2 are not too large, say at most R , one can hope to get a reasonable estimate for the coefficients $\#\{n \in \mathcal{A} : [d_1, d_2] | n\}$ of this

quadratic form. The best upper bound stemming from this method then comes from minimizing the quadratic form over all choices of $\lambda_d \in \mathbb{R}$ with $\lambda_1 = 1$ and $\lambda_d = 0$ for $d > R$.

For typical sets \mathcal{A} that arise in arithmetic problems, one finds that the optimal choice for the λ_d takes the form

$$\lambda_d \approx \mu(d) \cdot \left(\frac{\log(R/d)}{\log R} \right)^A \quad (d \leq R),$$

where A is some positive constant. We note that the weights λ_d decay to 0, and the larger the value of A , the higher the level of smoothness at the truncation point R . In the optimal choice, the exponent A is taken to be κ , the *dimension* of the sieve problem. However, for a given dimension κ , it is known [22, pg. 154] that any exponent $A > \kappa - 1/2$ yields weights λ_d whose dominant contribution comes from numbers almost coprime to m , whereas this fails to be true for smaller A . See [10, ch. 10] for further discussion.

More generally, one can consider the smoothed sieve weight

$$M_f(n; R) := \sum_{d|n} \mu(d) f\left(\frac{\log d}{\log R}\right),$$

where $f : \mathbb{R} \rightarrow \mathbb{R}$ is a function supported on $(-\infty, 1]$, which corresponds to taking $\lambda_d = \mu(d)f(\log d/\log R)$ for $d \leq R$. In Selberg sieve arguments one typically chooses f to be a polynomial in $[0, 1]$, perhaps of high degree. Such an example is offered by the ‘GPY sieve’ of Goldston-Pintz-Yıldırım [13, 27]. In more recent developments on gaps between primes by the third author [18] and Tao [24] one works with general smooth functions f .

The main motivation of this paper is to understand the exact role of the smoothing in the structure of the Selberg sieve weights. To this end, we consider their moments

$$\sum_{n \leq x} M_f(n; R)^k$$

as a tool of gaining additional insight on the distribution of the values of $M_f(n; R)$. On the practical side, higher moments naturally appear when applying Hölder’s inequality, so it would be useful to know their behaviour¹.

From the discussion above, in the case $f(x) = \max(1 - x, 0)^A$ and $k = 2$, we have seen that if A is sufficiently large, then $M_f(n; R)^2$ ‘behaves like a sieve weight’ in the sense that the sum $\sum_{n < x} M_f(n; R)^2$ is $O_f(x/\log R)$ and the main contribution to this comes from numbers with few prime factors less than R . If A is too small and so f is not smooth enough, however, then $M_f(n; R)$ exhibits qualitatively different behavior; the sum is larger than $x/\log R$, and the main contribution is no longer from numbers with few prime factors $\leq R$.

How smooth should f be so that $M_f(n; R)^{2k}$ behaves like a sieve weight when k varies, that is to say the main contribution to the $2k$ -th moment² of $M_f(n; R)$ comes from integers a that have very few prime factors $\leq R$? What happens in the extreme case where f is the discontinuous function $\mathbf{1}_{(-\infty, 1]}$? These are the types of questions that we will study in this paper.

¹For example, Lemma 3.5 in Pollack’s paper [23] is an example of a case where a fourth moment occurs because of the use of Cauchy’s inequality, and a similar issue is encountered in Friedlander’s work [9] for the combinatorial sieve instead of the Selberg sieve.

²We are typically interested in how large sieve weights get. If we took odd powers there might be an irrelevant cancellation, so we focus on even moments.

1.1. Some smoothing is necessary to behave like a sieve weight. In order to gain a first understanding of the importance of smoothing, let us consider the sharp cut-off function

$$f_0 := \mathbf{1}_{(-\infty, 1]}.$$

If $n = 2m$ with m odd, then we have that

$$(1.2) \quad M_{f_0}(n; R) = \sum_{\substack{d|n \\ d \leq R}} \mu(d) = \sum_{\substack{d|m \\ d \leq R}} \mu(d) + \sum_{\substack{d|m \\ 2d \leq R}} \mu(2d) = \sum_{\substack{d|m \\ R/2 < d \leq R}} \mu(d) = M_{\tilde{f}_0}(m; R),$$

where, with a slight abuse of notation, we have put

$$(1.3) \quad \tilde{f}_0 := \mathbf{1}_{(1 - \log 2 / \log R, 1]}.$$

In particular, if m is square-free and has exactly one divisor $d \in (R/2, R]$, then $M_{f_0}(n; R) = \pm 1$. An easy generalization of a deep result of Ford [7, Theorem 4] implies that³ the proportion of such $m \leq x/2$ is $\gg (\log R)^{-\delta} (\log \log R)^{-3/2}$ with

$$\delta = 1 - \frac{1 + \log \log 2}{\log 2} = 0.086071332 \dots,$$

whence we conclude that

$$\#\{n \leq x : M_{f_0}(n; R) \neq 0\} \gg \frac{x}{(\log R)^\delta (\log \log R)^{3/2}} \quad (x \geq R^{1+\epsilon}).$$

In particular, we find that $M_{f_0}(n; R)$ is non-zero too often to behave like a sieve weight. This indicates that part of the importance of smoothing is to reduce the contribution of isolated divisors of n to $M_f(n; R)$.

We will prove in Section 5 that

$$(1.4) \quad \#\{n \leq x : M_{f_0}(n; R) \neq 0\} \asymp \frac{x}{(\log R)^\delta (\log \log R)^{3/2}} \quad (x \geq R^{5/2}).$$

This sharpens a result by Hall and Tenenbaum [12], who used a very similar argument and the best results about divisors of integers available at that time.

1.2. A heuristic argument. Going back to the study of $M_f(n; R)$ for a smooth function f , it is reasonable to believe that the smoother f is, the larger the k are for which $M_f(n; R)^{2k}$ behaves like a sieve weight. One way to explain this phenomenon is by noticing that various integral transformations have faster decay for smooth weights, which can help to tame the arithmetic issues at play. (See, for example, Section 6.) Nevertheless, we prefer to give a number theoretic explanation in terms of the underlying sieve questions rather than an analytic one focused more on the technical issues. Assume that $n = p_1^{\alpha_1} \cdots p_r^{\alpha_r} m$, where $p_1 < \cdots < p_r$, $\alpha_i \geq 1$ and all of the prime divisors of m are $> p_r$. Then

$$(1.5) \quad \begin{aligned} M_f(n; R) &= \sum_{d|p_2 \cdots p_r m} \mu(d) f\left(\frac{\log d}{\log R}\right) + \sum_{d|p_2 \cdots p_r m} \mu(p_1 d) f\left(\frac{\log(p_1 d)}{\log R}\right) \\ &= \sum_{d|p_2 \cdots p_r m} \mu(d) \left\{ f\left(\frac{\log d}{\log R}\right) - f\left(\frac{\log p_1}{\log R} + \frac{\log d}{\log R}\right) \right\}. \end{aligned}$$

³The key estimates in the proof of the lower bound of Theorem 4 in [7] are the second part of Lemma 4.1, Lemma 4.3 (the parameters are $z = R \ll R/2 = y$), Lemma 4.5, Lemma 4.8 and Lemma 4.9. A key observation is that only square-free integers are considered in Lemma 4.8, so that a stronger version of the lower bound of Theorem 4 of [7] can be immediately deduced by the same proof, that counts square-free integers with exactly one divisor in $(R/2, R]$.

Continuing as above, we find that

$$(1.6) \quad M_f(n; R) = (-1)^r \sum_{d|m} \mu(d) \Delta^{(r)} f \left(\frac{\log d}{\log R}, \frac{\log p_1}{\log R}, \dots, \frac{\log p_r}{\log R} \right),$$

where $\Delta^{(r)} f(x; h_1, \dots, h_r)$ denotes the multi-difference operator defined by

$$\Delta^{(1)} f(x; h) = f(x + h) - f(x)$$

and

$$\Delta^{(r)} f(x; h_1, \dots, h_r) = \Delta^{(r-1)}(x + h_r; h_1, \dots, h_{r-1}) - \Delta^{(r-1)}(x; h_1, \dots, h_{r-1}).$$

In particular, if $f \in C^r(\mathbb{R})$, then

$$(1.7) \quad \Delta^{(r)} f(x; h_1, \dots, h_r) = \int_0^{h_r} \int_0^{h_{r-1}} \dots \int_0^{h_1} f^{(r)}(x + t_1 + t_2 + \dots + t_r) dt_1 \dots dt_r,$$

Returning to (1.6), we see that if $f \in C^A(\mathbb{R})$ and $n = p_1^{\alpha_1} \dots p_r^{\alpha_r} m$, $r \leq A$, is as above, then $M_f(n; R)$ should heuristically be $\ll M_{f^{(r)}}(m; R) \prod_{j=1}^r (\log p_j / \log R)$. Loosely, this indicates each additional degree of smoothness of the weight function f cuts the average size of $M_f(n; R)$ by about a factor of $1/\log R$.

The above discussion leads us to conjecture that if $f \in C^A(\mathbb{R})$ with $f(0) \neq 0$, then

$$(1.8) \quad \sum_{n \leq x} M_f(n; R)^{2k} \ll \max \left\{ \frac{x}{\log R}, \frac{1}{(\log R)^{2kA}} \sum_{n \leq x} M_{f_0}(n; R)^{2k} \right\}.$$

Notice that the factor $x/\log R$ is necessary because $M_f(n; R) = f(0)$ for all integers n that are free of prime factors $\leq R$.

Naturally, for this relation to be useful, we need to understand the asymptotics of $\sum_{n \leq x} M_{f_0}(n; R)^{2k}$. Recall the relation (1.9). Expanding the k -th power and swapping the order of summation, we find that

$$\sum_{n \leq x} M_f(n; R)^k = x \cdot \mathcal{M}_{f,k}(R) + O((\|f\|_\infty R)^k)$$

for any $f : \mathbb{R} \rightarrow \mathbb{R}$ supported on $(-\infty, 1]$, where

$$(1.9) \quad \mathcal{M}_{f,k}(R) := \sum_{d_1, \dots, d_k \geq 1} \frac{\prod_{j=1}^k \mu(d_j) f(\log d_j / \log R)}{[d_1, \dots, d_k]} = \prod_{p \leq R} \left(1 - \frac{1}{p}\right) \sum_{p|n \Rightarrow p \leq R} \frac{M_f(n; R)^k}{n}.$$

We are generally interested in the situation when R is bounded by a small power of x , so that the error term $O((\|f\|_\infty R)^k)$ is negligible. Thus our focus is on the main term $\mathcal{M}_{f,k}(R)$, which no longer depends on x . When $k = 1$, Dress, Iwaniec and Tenenbaum [4] showed

$$(1.10) \quad \mathcal{M}_{f_0,2}(R) \sim c_1 \quad (R \rightarrow \infty)$$

for some constant $c_1 > 0$, and when $k = 2$, Motohashi [19] showed that

$$(1.11) \quad \mathcal{M}_{f_0,4}(R) \sim c_2 (\log R)^2 \quad (R \rightarrow \infty)$$

for some constant $c_2 > 0$. In general, Balazard, Naimi, and Pétermann [1] proved that

$$\mathcal{M}_{f_0,2k}(R) = P_k(\log R) + O(e^{-c(\log R)^{3/5}(\log \log R)^{-1/5}}),$$

for some polynomial P_k and some constant $c = c(k) > 0$. This built on work of de la Bretèche [2], who showed how a wide class of related sums can be evaluated asymptotically. However, when

applying his technique to this question, one would need some strong understanding of the growth of $\zeta(s)$ near to $s = 1$ to recover the result of [1] (which, for example, follows from the Riemann Hypothesis).

Notice that if $\mathcal{E}_k = \deg(P_k)$, so that $\mathcal{E}_1 = 0$ and $\mathcal{E}_2 = 2$, then (1.8) becomes

$$(1.12) \quad \sum_{n \leq x} M_f(n; R)^{2k} \ll \max \left\{ \frac{x}{\log R}, x(\log R)^{\mathcal{E}_k - 2kA} \right\}.$$

This suggests that $M_f(n; R)^{2k}$ acts like a sieve weight as long as $A > \mathcal{E}_k/2k$. The big issue with the result of Balazard, Naimi and Pétermann is that the degree \mathcal{E}_k is not determined for general k , and that is essential if one wishes to gain a better understanding of how the Selberg sieve weights work. Our attention thus turns to calculating \mathcal{E}_k .

But first, we study seemingly analogous questions (in different settings), that one might guess would be easier and indicate what kind of estimate we should be looking for

1.3. Analogous settings. It is well-known that many of the analytic properties of integers are shared by both polynomials of finite fields (c.f. [20]), and by permutations (c.f. [11]). Moreover, polynomials and permutations are usually easier objects to understand, so in order to gain an understanding of the exponent \mathcal{E}_k , it would be natural to consider what happens in these analogous settings first.

Permutations. The easiest analogy to analyze concerns permutations. Every $\sigma \in S_N$ (the permutations on N letters) can be decomposed in a unique way into a product of disjoint cycles. Those cycles cannot be decomposed any further and play the role of irreducibles. Divisors of σ are precisely the set of possible products of cycles. If those cycles act on the subset T of $[N]$, then σ fixes T . Moreover, if σ fixes T , then σ is a product of cycles, a subproduct of which fixes T . Hence “divisors” correspond to sets $T \subset [N]$ for which $\sigma(T) = T$

To “calibrate” our understandings of the properties of integers and permutations, we note that for a typical integer n , its j -th largest prime factor is about e^{e^j} , whereas for a typical permutation $\sigma \in S_N$, its j -th largest cycle has length about e^j . Thus, the inequality $R/2 < d \leq R$ for a divisor d of n corresponds to having a set T that is fixed by σ of size $\#T = \log R + O(1)$. Hence we will study

$$(1.13) \quad \text{Perm}(N, m; k) := \frac{1}{N!} \sum_{\sigma \in S_N} \left| \sum_{\substack{T \subset [N], \#T=m, \\ \sigma(T)=T}} \mu(\sigma|_T) \right|^{2k},$$

where

$$[N] := \{1, \dots, N\},$$

and if $\sigma|_T = C_1 C_2 \dots C_\ell$ is the product of ℓ disjoint cycles, then we have set $\mu(\sigma|_T) = (-1)^\ell$. We claim that $\text{Perm}(N, m; k)$ is more natural than it appears at first sight. A usual function of permutations is the signature $\epsilon(\sigma)$ which counts the number of transpositions (i.e. the number of interchanges of two elements) needed to create σ . For a cycle C , one knows that $\epsilon(C) = (-1)^{\#C-1}$ and hence $\epsilon(\sigma|_T) = \epsilon(C_1)\epsilon(C_2)\dots\epsilon(C_\ell) = (-1)^{\#T-\ell} = (-1)^m \mu(\sigma|_T)$, since $\#T = m$ here. Therefore

$$\sum_{\substack{\sigma \in S_N, T \subset [N], \\ \#T=m, \sigma(T)=T}} \mu(\sigma|_T) = (-1)^m \sum_{\substack{\sigma \in S_N, T \subset [N], \\ \#T=m, \sigma(T)=T}} \epsilon(\sigma|_T),$$

whence

$$\text{Perm}(N, m; k) = \frac{1}{N!} \sum_{\sigma \in S_N} \left| \sum_{\substack{T \subset [N], \#T=m, \\ \sigma(T)=T}} \epsilon(\sigma|_T) \right|^{2k}.$$

Arguing as in the work of Eberhard, Ford and Green [5] that establishes the analogue for permutations of Ford's results [7] for integers, it is possible to show that the summands on the right hand of (1.13) (and, hence, of the above formula) are non-zero for a proportion $\asymp 1/m^\delta (\log m)^{3/2}$ of the permutations in S_N . The following theorem provides a formula and an asymptotic estimate for $\text{Perm}(N, m; k)$.

Theorem 1.1. *For each integer $k \geq 1$ and each integer $m \geq 1$, if $N \geq 2mk$ then*

$$\text{Perm}(N, m; k) = c(m, k),$$

where $c(m, k)$ is the number of $(2^{2k} - 1)$ -tuples $(r_I)_{\emptyset \neq I \subset \{1, \dots, 2k\}}$ of non-negative integers such that

- $r_I \in \{0, 1\}$ for $\#I$ odd;
- $\sum_{I: i \in I} r_I = m$, for each $i \in \{1, \dots, 2k\}$.

Moreover, for fixed $k \in \mathbb{Z}_{\geq 1}$, the function $c(m, k)$ is increasing in m and satisfies the estimate

$$c(m, k) \asymp_k m^{2^{2k-1} - 2k - 1} + 1.$$

Proof of the formula for $\text{Perm}(N, m; k)$. Given sets T_1, \dots, T_{2k} , the sets

$$R_I := \left(\bigcap_{i \in I} T_i \right) \setminus \left(\bigcup_{i \in [2k] \setminus I} T_i \right) \quad (I \subset [2k])$$

form a partition of $[N]$; that is to say $[N]$ equals $\sqcup_I R_I$, the disjoint union of the sets R_I . Using this with T_1, \dots, T_{2k} fixed sets of σ (i.e. $\sigma(T_i) = T_i$, so the R_I are all fixed by σ as well), we find

$$\frac{1}{N!} \sum_{\sigma \in S_N} \left| \sum_{\substack{T \subset [N], \#T=m, \\ \sigma(T)=T}} \epsilon(\sigma|_T) \right|^{2k} = \sum_{\substack{r_I \geq 0 \forall I \\ \sum_{I: i \in I} r_I = m}} \sum_{\substack{[N] = \sqcup_I R_I \\ \#R_I = r_I \forall I}} \frac{1}{N!} \prod_{I \subset [2k]} \left(\sum_{\rho_I \in S_{r_I}} \epsilon(\rho_I)^{\#I} \right).$$

The inner sums are each $r_I!$ unless $\#I$ is odd and $r_I > 1$, in which case we get 0. Additionally, we get that the number of choices of sets of the given sizes is $N! / \prod_I r_I!$, and hence the above equals $c(m, k)$.

The bounds for $c(m, k)$ will be proven in Section 3. \square

Evidently, the above results suggest that $\mathcal{E}_k = \max\{0, 2^{2k-1} - 2k - 1\}$. Relation (1.10) implies that $\mathcal{E}_1 = 0$, but relation (1.11) implies that $\mathcal{E}_2 = 2 \neq 2^3 - 5$. This suggests that there is a discrepancy between the integer and the permutation setting, a very rare exception.

Polynomials over finite fields. Positive integers are uniquely identifiable by their factorization into primes (the Fundamental Theorem of Arithmetic). Note though that every non-zero integer equals a unit (that is 1 or -1) times one of those positive integers. We will work with polynomials in $\mathbb{F}_q[t]$. Monic polynomials in $\mathbb{F}_q[t]$ are uniquely identifiable by their factorization into monic irreducible polynomials of degree ≥ 1 . Again, note that every non-zero polynomial in $\mathbb{F}_q[t]$ equals a unit (that is, any element $a \in \mathbb{F}_q \setminus \{0\}$) times a monic polynomial. We will work only with monic polynomials, for example when considering divisors of a given polynomial (rather like we only consider positive integer divisors of a given integer). The Möbius function of a given polynomial is a multiplicative function, where $\mu(P) = -1$, and $\mu(P^k) = 0$ if $k \geq 2$, whenever P is irreducible.

To “calibrate” our understandings of the arithmetic properties of integers and polynomials, we note that $\sim 1/\log x$ of integers around x are prime, whereas $\sim 1/m$ of monic polynomials of degree m are irreducible in $\mathbb{F}_q[t]$. Here the “ \sim ” symbol means as $q \rightarrow \infty$ running through prime powers. Thus, wherever we see $\log x$ in an estimate about the integers, we try to replace it with m in an estimate about degree m polynomials. Similarly a divisor d of n that is close to R is analogous to a polynomial divisor of $F(t)$ of degree m , where m replaces $\log R$ in estimates. Hence we will study

$$\text{Poly}_q(n, m; k) := \frac{1}{q^n} \sum_{\substack{\text{monic } N \in \mathbb{F}_q[t] \\ \deg N = n}} \left| \sum_{\substack{\text{monic } M | N \\ \deg M = m}} \mu(M) \right|^{2k},$$

Here we have divided by q^n because this is how many monic polynomials N of degree n are contained in $\mathbb{F}_q[t]$, which is the analogue of

$$\frac{1}{x} \sum_{n \leq x} \left(\sum_{\substack{d | n \\ R/2 < d \leq R}} \mu(d) \right)^{2k},$$

a quantity directly related to $\frac{1}{x} \sum_{n \leq x} M_{f_0}(n; R)^{2k}$ via (1.2). We will prove below the following estimate:

Theorem 1.2. *For integers $k, m \geq 1$ and $n \geq 2mk$, we have that*

$$\text{Poly}_q(n, m; k) = c(m, k)(1 + O_k(1/q)) \asymp_k 1 + m^{2^{2k-1}-2k-1}.$$

We thus see that polynomials behave similarly to permutations (and thus differently than integers).

1.4. Two worlds apart and a bridge between them. Our discussion of the permutation and polynomial analogues, rather than shedding more light on the value of the exponent \mathcal{E}_k , gave rise to even more questions. It turns out that the integer setting is substantially more complicated than the permutation and polynomial settings. We now state our main results about integers. First, given $A \in \mathbb{Z}_{\geq 1}$, we let

$$f_A(t) := \begin{cases} (1-t)^A & \text{for } t \leq 1; \\ 0 & \text{otherwise,} \end{cases}$$

an extension of the definition of f_0 . Note that $f_A \in C^{A-1}(\mathbb{R}) \setminus C^A(\mathbb{R})$ for all $A \geq 1$. We then have the following result that determines the value of \mathcal{E}_k :

Theorem 1.3. *For fixed integers $k \geq 1$ and $A \geq 0$, there is a constant $c_{k,A} > 0$ such that*

$$(1.14) \quad \mathcal{M}_{f_A, 2k}(R) = c_{k,A}(\log R)^{\mathcal{E}_{k,A}} + O((\log R)^{\mathcal{E}_{k,A}-1}),$$

where

$$\mathcal{E}_{k,A} := \max \left\{ \binom{2k}{k} - 2k(A+1), -1 \right\}.$$

In particular, $\mathcal{E}_k = \mathcal{E}_{k,0} = \binom{2k}{k} - 2k$. Additionally, we find that there is a constant $c'_k > 0$ such that for $R^{2k} \leq x$ we have

$$(1.15) \quad \frac{1}{x} \sum_{n \leq x} \left(\sum_{\substack{d | n \\ R/2 < d \leq R}} \mu(d) \right)^{2k} = c'_k (\log R)^{\binom{2k}{k} - 2k} + O \left((\log R)^{\binom{2k}{k} - 2k - 1} \right).$$

All implied constants depend at most on k and A .

Remark 1.1. We have no nice formula for the constants $c_{k,A}$ and c'_k appearing in Theorem 1.3; we only know how to write them as an enormous rational linear combination of complicated integrals, and leave it as a challenge to come up with an easy explicit description.

Remark 1.2. If the moments of a distribution grow slowly, then the distribution can be determined via its Laplace transform. However, in our case the moments are of rapidly increasing magnitude, indeed with different powers of $\log R$, so one cannot immediately deduce from them the distribution of the weights $M_{f_A}(n; R)$ as n varies over the integers.

Remark 1.3. In this paper we only consider integral A , but we would expect analogous results to hold for all real $A > 0$.

Remark 1.4. In this paper we only consider Selberg-style sieve weights. We would expect something somewhat analogous to hold for combinatorial-style sieve weights (such as those used in the β -sieve) but we do not consider such situations here.

For general functions f , we prove that $M_f(n; R)^{2k}$ behaves like a sieve weight as long $f \in C^A(\mathbb{R})$ with $A > \binom{2k}{k}/2k = \mathcal{E}_{2k}/2k + 1$. Notice that this confirms a weak version of the heuristic estimate (1.12).

Theorem 1.4. *Let $k \in \mathbb{Z}_{\geq 1}$, $\epsilon \in (0, 1)$ and $f : \mathbb{R} \rightarrow \mathbb{R}$ be supported in $(-\infty, 1]$. Assume further that for some integer $A \geq 2$, $f \in C^A(\mathbb{R})$ and that all functions $f, f', \dots, f^{(A)}$ are bounded.*

(a) *If $A > \frac{1}{2k} \binom{2k}{k}$, then for $x \geq R \geq 2$ and $1 \geq \eta \geq \log 2 / \log R$, we have that*

$$\sum_{\substack{n \leq x \\ \exists p|n, p \leq R^\eta}} M_f(n; R)^{2k} \ll \frac{\eta x}{\log R}.$$

If, in addition, $f(0) \neq 0$, then there is a constant $c_{k,f} > 0$ such that for $x \geq R^{2k} \log^2 R$ we have that

$$\frac{1}{x} \sum_{n < x} M_f(n; R)^{2k} = \frac{c_{k,f}}{\log R} + O\left(\frac{1}{(\log R)^{2-\epsilon}}\right).$$

(b) *If $A \leq \frac{1}{2k} \binom{2k}{k}$, then for $x \geq R \geq 2$ we have that*

$$\sum_{n \leq x} M_f(n; R)^{2k} \ll x(\log R)^{\binom{2k}{k} - 2kA}.$$

All implied constants depend at most on f, k and ϵ .

The value of $\mathcal{E}_k = \binom{2k}{k} - 2k$ given by Theorem 1.3 is significantly smaller than the exponent $2^{2k-1} - 2k - 1$ in the polynomial/permutation setting. So we see the usual analogy breaking down in quite a severe way, something surprising. We devote Section 2 to the analysis of this discrepancy. In particular, we will see that the underlying reason is the relation

$$(1.16) \quad \sum_{\substack{\sigma \in S_N, T \subset [N], \\ \#T=m, \sigma(T)=T}} \mu(\sigma|_T) = (-1)^m \sum_{\substack{\sigma \in S_N, T \subset [N], \\ \#T=m, \sigma(T)=T}} \epsilon(\sigma|_T)$$

that we saw before. Notice here that while $\mu(\rho) = -1$ for all cycles ρ , we have that $\epsilon(\rho)$ takes the values ± 1 with equal probability as ρ ranges over cycles of all possible lengths. The simplest

example of a multiplicative function over \mathbb{Z} demonstrating this kind of behaviour is that of a real Dirichlet character. To this end, we consider

$$\mathcal{X}_{2k}(R) = \prod_{p \leq R} \left(1 - \frac{1}{p}\right) \sum_{P^+(n) \leq R} \frac{1}{n} \left(\sum_{\substack{d|n \\ R/2 < d \leq R}} \chi(d) \right)^{2k},$$

which, as in (1.9), is the main term of

$$\frac{1}{x} \sum_{n < x} \left| \sum_{\substack{d|n \\ R/2 < d \leq R}} \chi(d) \right|^{2k}.$$

We then have the following theorem, which shows that it is possible to bridge the gap between the two worlds of integers and of permutations/polynomials. All implied constants below depend at most on k , and we have set

$$\mathcal{S}^+(2k) := \{I \subset \{1, 2, \dots, 2k\} : \#I \text{ even}\} \setminus \{\emptyset\}.$$

Theorem 1.5. *Let $\chi \pmod{q}$ be a real non-principal character and $k \in \mathbb{Z}_{\geq 1}$.*

(a) *If $k = 1$, then*

$$\mathcal{X}_2(R) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{P(t, \chi) |L(1 + it, \chi)|^2 \sin^2(t(\log 2)/2)}{t^2} dt + O\left(\frac{1}{(\log R)^{100}}\right),$$

where $P(\cdot, \chi)$ is a real-valued Euler product whose factors are $1 + O(1/p^2)$. In particular, $P(t, \chi) \asymp 1$ for all t , uniformly in χ .

(b) *Assume that $k \geq 2$. Let $V_k(m)$ be the Lebesgue volume in $\mathbb{R}^{2^{2k-1}-1}$ given by*

$$V_k(m) = \text{vol}\{(x_I)_{I \in \mathcal{S}^+(2k)} : x_I \geq 0, m - \log 2 \leq \sum_{I \ni i} x_I \leq m\},$$

and let $\mathfrak{S}_k(\chi)$ be the singular series

$$\mathfrak{S}_k(\chi) = \prod_p \left(1 - \frac{1}{p}\right)^{2^{2k-1}} f_p,$$

where

$$f_p = \begin{cases} \sum_{j \geq 0} (j+1)^{2k}/p^j, & \text{if } \chi(p) = 1, \\ (1 - 1/p^2)^{-1}, & \text{if } \chi(p) = -1, \\ (1 - 1/p)^{-1}, & \text{if } p|q. \end{cases}$$

Then $V_k(m) \asymp_k m^{2^{2k-1}-2k-1}$, $\mathfrak{S}_k(\chi) \asymp_k L(1, \chi)^{2^{2k-1}}$, and

$$\mathcal{X}_{2k}(R) = \mathfrak{S}_k(\chi) \cdot V_k(\log R) + O\left((\log R)^{2^{2k-1}-2k-2} (\log(q \log R))^{O(1)}\right).$$

(c) *Assume that $k \geq 2$ and that $L(\beta, \chi) = 0$ for some $\beta > 1 - 1/(100 \log q)$. If $Q = e^{1/(1-\beta)}$ and $e^{(\log q)^C} \leq R \leq Q$ for some large enough $C = C(k)$, then there is a constant $c_k(\chi) = (\log q)^{O(1)}$ such that*

$$\mathcal{X}_{2k}(R) = c_k(\chi) (\log R)^{\binom{2k}{k} - 2k} \left(1 + O\left(\frac{(\log(q \log R))^{O(1)}}{\log R}\right)\right).$$

In the case of our polynomial and permutation models, we have an exponent of $2^{2k-1} - 2k - 1$ for the $2k^{\text{th}}$ moment with $k \geq 2$, whilst over the integers we have an exponent $\binom{2k}{k} - 2k$. We see that our Dirichlet character model interpolates between these two settings. If the Dirichlet L -function associated with the character has a zero very close to 1, then $\chi(p) = -1$ for many small primes p , and so by multiplicativity χ behaves similarly to μ (at least in appropriate ranges). This is represented by our exponent of $\binom{2k}{k} - 2k$ in this case. On the other hand, χ is a periodic character, and if the L function does not have a zero very close to 1, we see that we have an exponent $2^{2k-1} - 2k - 1$, matching the exponent of our polynomial and permutation models. Notice that if $L(s, \chi)$ does have an exceptional zero, then the asymptotic of case (c) for $\mathcal{X}_{2k}(R)$ holds for small R , and transitions into the asymptotic of case (b) as R grows.

Remark 1.5. Relation (1.16) has a polynomial analogue whose consequences are worth exploring further. Given $I \subset \{F \in \mathbb{F}_q[t] : \deg(F) = n\}$, we consider the sum

$$\sum_{F \in I} \mu(F).$$

For example, we could take $I = \{F \in \mathbb{F}_q[t] : \deg(F) = n\}$, or $I = \{F \in \mathbb{F}_q[t] : \deg(F - F_0) \leq h\}$ for some $F_0 \in \mathbb{F}_q[t]$ of degree n and for some integer $h \in [1, n-1]$, which can be seen as the polynomial analogue of a short interval. Then

$$\sum_{F \in I} \mu(F) = (-1)^n \sum_{F \in I} \chi(F),$$

where $\chi(F) = (-1)^{\deg(F)} \mu(F)$, which is also a multiplicative function. However, we note that, even though $\mu(P) = -1$ for all irreducibles, we have that $\chi(P) = 1$ for about half of the irreducibles P , and $\chi(P) = -1$ for the other half, that is to say χ behaves on average much more like a real Dirichlet character rather than the Möbius function.

This phenomenon is striking and sharply different than what happens over \mathbb{Z} , where there is a dichotomy between multiplicative functions that look like the Möbius functions and other ones whose average prime value is 0, as is exemplified by Theorem 1.5 (see, also, [17]).

1.5. Further analysis of truncated Möbius divisor sums. As we saw in Theorem 1.4, if $f \in C^A(\mathbb{R})$ with $A > \frac{1}{2k} \binom{2k}{k} = \mathcal{E}_k/2k + 1$, then $M_f(n; R)^{2k}$ behaves like a sieve weight. When $f = f_A$, we can be more precise:

If $A > \frac{1}{2k} \binom{2k}{k} - 1 = \mathcal{E}_k/2k$, then $\mathcal{E}_{k,A} = -1$ in Theorem 1.3, and so $\mathcal{M}_{f_A, 2k}(R) \asymp (\log R)^{-1}$. Since integers $n \leq x$ with no prime factors less than R contribute a total $\gg x(\log R)^{-1}$ to $\sum_{n \leq x} M_{f_A}(n; R)^{2k}$, we see that $M_{f_A}(n; R)^{2k}$ is behaving like a sieve weight in this case.

If $A \leq \frac{1}{2k} \binom{2k}{k} - 1 = \mathcal{E}_k/2k$, then $\mathcal{E}_{k,A} \geq 0$, and so $\mathcal{M}_{f_A, 2k}(R) \gg 1$. In particular, $M_{f_A}(n; R)^{2k}$ no longer behaves like a sieve weight, and the main contribution is from numbers with several prime factors in $[1, R]$.

The following theorem illustrates further this distinction.

Theorem 1.6. *Let $x \geq R \geq 2$, $k \in \mathbb{Z}_{\geq 1}$ and $A \in \mathbb{Z}_{\geq 0}$. Moreover, let $\Omega(n; R)$ denote the number of prime factors of n in $[1, R]$, counted with multiplicity.*

(a) *If $A > \frac{1}{2k} \binom{2k}{k} - 1$, then*

$$\sum_{\substack{n \leq x \\ \Omega(n; R) \geq C}} M_{f_A}(n; R)^{2k} \ll_{k,A} \frac{x}{C \log R}.$$

(b) If $A \leq \frac{1}{2k} \binom{2k}{k} - 1$ and $\epsilon > 0$ is fixed, then there is a $\delta = \delta(\epsilon, k) > 0$ such that

$$\sum_{\substack{n < x \\ |\Omega(n; R) / \log \log R - \binom{2k}{k}| \geq \epsilon}} M_{f_A}(n; R)^{2k} \ll_{k,A} x (\log R)^{\binom{2k}{k} - 2k(A+1) - \delta}.$$

In other words, if $A > \frac{1}{2k} \binom{2k}{k} - 1$, then the main contribution to the sum defining $\mathcal{M}_{f_A, 2k}(R)$ comes from integers with a bounded number of prime factors $\leq R$; whereas if $A \leq \frac{1}{2k} \binom{2k}{k} - 1$, then the main contribution to the sum comes from integers with $(\binom{2k}{k} + o(1)) \log \log R$ prime factors $\leq R$.

Analogous results hold with $\Omega(n; R)$ replaced by the function $\#\{p|n : p \leq R\}$. We note that typically one requires $x > R^{2k}$, as in Theorem 1.3, to estimate a $2k^{\text{th}}$ moment of a sum of divisors of size at most R , but the estimates of Theorem 1.6 hold in the much wider range $x \geq R$. We can show similar (but slightly weaker) results for general weights f :

Theorem 1.7. *Let $k \in \mathbb{Z}_{\geq 1}$ and $f : \mathbb{R} \rightarrow \mathbb{R}$ be supported in $(-\infty, 1]$. Assume further that $f \in C^A(\mathbb{R})$ and that all functions $f, f', \dots, f^{(A)}$ are uniformly bounded for some integer $A \geq 2$, and fix some $\epsilon \in (0, 1)$.*

(a) Assume that $A > \frac{1}{2k} \binom{2k}{k}$. For $x \geq R \geq 2$ and $C \geq 1$, we have that

$$\sum_{\substack{n < x \\ \Omega(n; R) \geq C}} M_f(n; R)^{2k} \ll_{k,f} \frac{x}{C \log R}.$$

(b) If $A \leq \frac{1}{2k} \binom{2k}{k}$ and $\epsilon > 0$ is fixed, then there is a $\delta = \delta(\epsilon, k) > 0$ such that

$$\sum_{\substack{n < x \\ |\Omega(n; R) / \log \log R - \binom{2k}{k}| \geq \epsilon}} M_f(n; R)^{2k} \ll_{k,f,\epsilon} x (\log R)^{\binom{2k}{k} - 2kA - \delta} \quad (x \geq R \geq 2).$$

1.6. Outline of the paper. We start the paper in Section 2 with a discussion of the discrepancy between the exponent of $\log R$ in (1.15) and the exponent of m in Theorem 1.1, which is surprising at first sight.

Sections 3 and 4 study the analogies for permutations and polynomials over finite fields, respectively. These analogies are considerably easier to analyze than the integer case.

The main body of the paper is then dedicated to the study of moments of $M_{f_A}(n; R)$ over Sections 5 – 10. Specifically, in Section 5 we establish relation (1.4) for the size of the support of $M_{f_0}(n; R)$, and in Section 6 we study inversion formulas for our divisor sums $M_f(n; R)$ that will be essential when dealing with their moments. The proof of Theorem 1.3 is separated over three sections: in Section 7, we establish certain combinatorial inequalities that will be instrumental in understanding the leading term in the asymptotics for $\mathcal{M}_{f_A, 2k}(R)$. Then, in Section 8 we establish Theorem 1.3 by a multidimensional contour shifting argument, except for showing the positivity of the constants $c_{k,A}$ and c'_k . The latter will be accomplished with a different argument in Section 9. Section 10 contains an analysis of the anatomy of the integers that give the main contribution to moments of $M_f(n; R)$. Specifically, we prove Theorems 1.4, 1.6 and 1.7 there.

Finally, in Sections 11 and 12 we study the moments of the sum weighted by Dirichlet characters, and establish Theorem 1.5, first for non-exceptional Dirichlet characters (where the proof is similar to Theorem 1.2), and then for exceptional Dirichlet characters (where the proof is similar to Theorem 1.3).

1.7. **Notation.** Given an integer $N \geq 1$, we set throughout the paper

$$[N] := \{1, 2, \dots, N\},$$

$$\mathcal{S}^+(N) := \{\emptyset \neq I \subset [N] : \#I \text{ even}\}, \quad \mathcal{S}^-(N) := \{I \subset [N] : \#I \text{ odd}\},$$

$$\mathcal{S}(N) := \{I \subset [N]\} \quad \text{and} \quad \mathcal{S}^*(N) := \mathcal{S}^+(N) \cup \mathcal{S}^-(N).$$

Also, we recall that, given a integer $n \geq 1$, we write $P^+(n)$ and $P^-(n)$ for its largest and smallest prime divisors, respectively, with the convention that $P^+(1) = 1$ and $P^-(1) = \infty$.

Finally, given $2k$ variables s_1, \dots, s_{2k} and $I \subset [2k]$, we will use the notation $s_I = \sum_{i \in I} s_i$.

2. THE DISCREPANCY BETWEEN INTEGERS AND POLYNOMIALS

The goal of this section is to analyze in detail why we have such a different behaviour when considering integers vs. polynomials or permutations

2.1. **Integer setting.** Assume that $k \geq 2$. We mimic the proof of Theorem 1.1. Recall the definition of \tilde{f}_0 in (1.3). Given square-free integers d_1, \dots, d_{2k} and $I \subset \mathcal{S}^*(2k)$, we let D_I be the product of those primes p that divide each of the d_i 's with $i \in I$ but do not divide $\prod_{i \in [2k] \setminus I} d_i$. Then the integers D_I for $I \in \mathcal{S}^*(2k)$ are pairwise coprime and $d_i = \prod_{I: i \in I} D_I$ for each i , so that

$$\begin{aligned} \mathcal{M}_{\tilde{f}_0, 2k}(R) &= \sum_{R/2 < d_1, \dots, d_{2k} < R} \frac{\mu(d_1) \dots \mu(d_{2k})}{[d_1, \dots, d_{2k}]} \\ &= \sum_{\substack{D_I \text{ (} I \in \mathcal{S}^*(2k) \text{)} \\ R/2 < \prod_{I \ni i} D_I \leq R \text{ (} 1 \leq i \leq 2k \text{)}}}^b \left(\prod_{I \in \mathcal{S}^-(2k)} \frac{\mu(D_I)}{D_I} \right) \left(\prod_{I \in \mathcal{S}^+(2k)} \frac{1}{D_I} \right), \end{aligned}$$

where the notation \sum^b means that the summation is running over squarefree and pairwise coprime integers D_I . Set $L = e^{(\log \log R)^3}$. The contribution of those tuples with $D_I > L$ for some I odd to $\mathcal{M}_{\tilde{f}_0, 2k}(R)$ can be seen to be $\ll 1/e^{c(\log \log R)^{3/2}}$ for some $c = c(k) > 0$, by the Prime Number Theorem. So assume that $D_I \leq L$ for all I odd. Then it is natural to write

$$\mathcal{M}_{\tilde{f}_0, 2k}(R) \sim \sum_{\substack{D_I \leq L \text{ (} I \in \mathcal{S}^*(2k) \text{)} \\ D = \prod_{I \in \mathcal{S}^-(2k)} D_I}}^b \frac{\mu(D)}{D} \cdot T_{2k}(R_1, \dots, R_{2k}; D),$$

where $R_i = R / \prod_{I \ni i, I \in \mathcal{S}^+(2k)} D_I$ and

$$T_{2k}(\mathbf{R}; a) = \sum_{\substack{(D_I, a)=1 \text{ (} I \in \mathcal{S}^+(2k) \text{)} \\ R_i/2 < \prod_{I \in \mathcal{S}^+(2k): i \in I} D_I \leq R_i \text{ (} 1 \leq i \leq 2k \text{)}}}^b \prod_{I \in \mathcal{S}^+(2k)} \frac{1}{D_I}.$$

When $\log R_i = \log R + O(\log L) = \log R + O((\log \log R)^{3/2})$, as it is here, we should be expecting that $T_{2k}(\mathbf{R}; a)$ has an asymptotic formula of the form

$$T_{2k}(\mathbf{R}; a) = g(a)(\log R)^{2^{2k-1}-2k-1} + O\left((\log R)^{2^{2k-1}-2k-2}(\log \log R)^{O(1)}\right),$$

since we have $2^{2k-1} - 1$ variables on a logarithmic scale and $2k$ multiplicative constraints in dyadic intervals, where $g(a)$ is a multiplicative function with $g(p) = 1 + O(1/p)$. Since $\sum_{n=1}^{\infty} \mu(n)/n = 0$, we then find that the total contribution of the main terms to $\mathcal{M}_{\tilde{f}_0, 2k}(R)$ is

$$(\log R)^{2^{2k-1}-2k-1} \sum_{\substack{D_I \leq L \\ D = \prod_{I \in \mathcal{S}^-(2k)} D_I}}^b \frac{\mu(D)g(D)}{D} \ll e^{-c'(\log \log R)^{3/2}}.$$

for some $c' = c'(k) > 0$, which is negligible. Consequently,

$$\mathcal{M}_{\tilde{f}_0, 2k}(R) \ll (\log \log R)^{O(1)} (\log R)^{2^{2k-1}-2k-2},$$

whereas the power of m in Theorem 1.1 is $2^{2k-1} - 2k - 1$. So this heuristic indicates that we should get more cancellation in the integer setting than we will obtain in the analogous permutation question, as established in Theorem 1.1.

2.2. Polynomial analogue. The reader might be sceptical of the argument presented above, because a direct analogue exists for polynomials over finite fields too. Specifically, expanding the $2k$ -th power in $\text{Poly}_q(n, m; k)$, we find that

$$\text{Poly}_q(n, m; k) = \sum_{G_1, \dots, G_{2k}} \frac{\mu(G_1) \cdots \mu(G_{2k})}{q^{\deg([G_1, \dots, G_{2k}])}}$$

for $n \geq 2mk$. Given square-free, monic polynomials G_1, \dots, G_{2k} over $\mathbb{F}_q[t]$ and $I \subset \mathcal{S}^*(2k)$, we let G_I be the product of those monic irreducibles P that divide each of the G_i 's with $i \in I$ but do not divide $\prod_{i \in [2k] \setminus I} G_i$. Then the polynomials G_I for $I \in \mathcal{S}^*(2k)$ are pairwise coprime and $G_i = \prod_{I: i \in I} G_I$ for each i , so that

$$\text{Poly}_q(n, m; k) = \sum_{\substack{G_I \ (I \in \mathcal{S}^*(2k)) \\ \sum_{I \ni i} \deg(G_I) = m \ (1 \leq i \leq 2k)}}^b \left(\prod_{I \in \mathcal{S}^-(2k)} \frac{\mu(G_I)}{q^{\deg(G_I)}} \right) \left(\prod_{I \in \mathcal{S}^+(2k)} \frac{1}{q^{\deg(G_I)}} \right).$$

where the notation \sum^b means that the summation is running over squarefree and pairwise monic polynomials G_I . As in the integer case, the contribution to $\text{Poly}_q(n, m; k)$ of those tuples $(G_I)_{I \in \mathcal{S}^*(2k)}$ such that $\deg(G_I)$ is large for some $I \in \mathcal{S}^-(2k)$ is negligible, by the Prime Number Theorem over $\mathbb{F}_q[t]$. Hence, we may assume that $\deg(G_I) \leq \log m$ for all $I \in \mathcal{S}^-(2k)$. Then it is natural to write

$$\text{Poly}_q(n, m; k) = \sum_{\substack{\deg(G_I) \leq \log m \ (I \in \mathcal{S}^-(2k)) \\ G = \prod_{I \in \mathcal{S}^-(2k)} G_I}}^b \frac{\mu(G)}{q^{\deg(G)}} \cdot \tilde{T}_{q, 2k}(m_1, \dots, m_{2k}; G),$$

where $m_i = m - \sum_{I \ni i, I \in \mathcal{S}^+(2k)} \deg(G_I)$ and

$$\tilde{T}_{q, 2k}(\mathbf{m}; A) := \sum_{\substack{(G_I, A) = 1 \ (I \in \mathcal{S}^+(2k)) \\ \sum_{I \in \mathcal{S}^+(2k): i \in I} \deg(G_I) = m_i \ (1 \leq i \leq 2k)}}^b \prod_{I \in \mathcal{S}^+(2k)} \frac{1}{q^{\deg(G_I)}}.$$

As before, when $\ell_i = m + O(\log m)$, as above, we should be expecting that $\tilde{T}_{q, 2k}(\ell; A)$ has an asymptotic formula of the form

$$(2.1) \quad \tilde{T}_{q, 2k}(\ell; A) = \tilde{g}(A) m^{2^{2k-1}-2k-1} + O\left(m^{2^{2k-1}-2k-2} (\log m)^{O(1)}\right),$$

where $\tilde{g}(A)$ is a multiplicative function with $\tilde{g}(P) = 1 + O(1/q^{\deg(P)})$ for irreducibles P .

The above argument suggests that we should have an asymptotic behaviour of $\text{Poly}_q(n, m; k)$ that is smaller than what Theorem 1.2 states, which is absurd. The problem is that if $\sum_{I \ni i} \deg(G_I) = m$ for all i , then we also have that

$$2km = \sum_{i=1}^{2k} \sum_{I: i \in I} \deg(G_I) = \sum_I \#I \deg(G_I).$$

Reducing this formula mod 2, we find that

$$(2.2) \quad \sum_{I \in \mathcal{S}^-(2k)} \deg(G_I) \equiv 0 \pmod{2},$$

a local constraint that is not present in the integer analogue. In particular, we see that (2.1) is true only when (2.2) is satisfied. We thus find that the main term for $\text{Poly}_q(n, m; k)$ equals

$$\begin{aligned} & m^{2^{2k-1}-2k-1} \cdot \sum_{\substack{\deg(G_I) \leq \log m \ (I \in \mathcal{S}^-(2k)) \\ G = \prod_{I \in \mathcal{S}^-(2k)} G_I, \ 2 \mid \deg(G)}}^b \frac{\mu(G) \tilde{g}(G)}{q^{\deg(G)}} \\ &= \frac{m^{2^{2k-1}-2k-1}}{2} \cdot \sum_{\substack{\deg(G_I) \leq \log m \ (I \in \mathcal{S}^-(2k)) \\ G = \prod_{I \in \mathcal{S}^-(2k)} G_I}}^b \frac{\mu(G)(1 + (-1)^{\deg(G)})}{q^{\deg(G)}} + O\left(\frac{1}{q}\right) \\ &\asymp m^{2^{2k-1}-2k-1}, \end{aligned}$$

because $\sum_F \mu(F)/q^{\deg(F)} = 0$ and $\sum_F \mu(F)(-1/q)^{\deg(F)} = \prod_P (1 - (-1/q)^{\deg(P)}) > 0$. Thus we see the local constraint associated to the discreteness of degrees in the polynomial setting means we have genuinely different asymptotic behavior.

2.3. Further analysis. The above arguments suggest a possible route to proving Theorem 1.3, by working out the full asymptotic expansion of $T_{2k}(\mathbf{x}; a)$. Controlling the coefficients in this expansion is a highly non-trivial problem. Instead, we take another route, using a high-dimensional contour shifting argument. Our starting point is Perron's inversion formula which, ignoring convergence issues, yields

$$T_{2k}(\mathbf{x}; a) \sim \frac{1}{(2\pi i)^{2k}} \int \cdots \int_{\substack{\text{Re}(s_j)=1/\log R \\ 1 \leq j \leq 2k}} \sum_{\substack{(D_I, a)=1 \\ I \in \mathcal{S}^+(2k)}}^b \prod_{I \in \mathcal{S}^+(2k)} \frac{1}{D_I^{1+s_I}} \prod_{j=1}^{2k} \frac{x_j^{s_j} (1 - 2^{-s_j})}{s_j} ds_1 \cdots ds_{2k},$$

with the notational convention that $s_I = \sum_{j \in I} s_j$. Therefore

$$\mathcal{M}_{\tilde{f}_0, 2k}(R) \sim \frac{1}{(2\pi i)^{2k}} \int \cdots \int_{\substack{\text{Re}(s_j)=1/\log R \\ 1 \leq j \leq 2k}} F(\mathbf{s}) \frac{\prod_{I \in \mathcal{S}^+(2k)} \zeta(1 + s_I)}{\prod_{I \in \mathcal{S}^-(2k)} \zeta(1 + s_I)} \prod_{j=1}^{2k} \frac{1 - 2^{-s_j}}{s_j} ds_1 \cdots ds_{2k},$$

where $F(\mathbf{s})$ is analytic and non-zero when $\text{Re}(s_j) > -1/4k$ for all j . As we will see in Section 8, shifting contours, we pick up poles any time $s_I = 0$ for some $I \in \mathcal{S}^+(2k)$. What is the difficulty in proving Theorem 1.3 is that some of these poles can get annihilated by poles of the zeta factors in the denominator, which is an analytic way of saying that the higher order terms in the asymptotic expansions of $T_{2k}(\mathbf{x}; a)$ are cancelled out.

It is clear from the above discussion that the underlying reason why we got a genuinely smaller main term for $\mathcal{M}_{\tilde{f}_0, 2k}(R)$ is the identity $\sum_{n=1}^{\infty} \mu(n)/n = 0$, that is to say the fact that $1/\zeta$ has a zero at 1. This also explains the phenomenon we see in Theorem 1.5. If we replace μ by a real valued multiplicative function f whose Dirichlet series $F(s) = \sum_{n=1}^{\infty} f(n)/n^s$ which is not very small at $s = 1$, then the behaviour of the respective divisor sums should be similar as the permutation analogue, whilst if $F(s)$ is close to 0 at $s = 1$ (which occurs if F has a zero very close to 1) the behaviour is the same as in the original integer setting.

2.4. Further obstructions to the analogy. Is it possible that the local constraints at the prime 2 described above are the only thing separating integers and polynomials? In order to study this question, we consider the variations

$$\text{Poly}_q(n, m, h; k) := \frac{1}{q^n} \sum_{\substack{N \in \mathbb{F}_q[t] \\ \deg N = n}} \left| \sum_{\substack{M|N \\ m-h < \deg M \leq m}} \mu(M) \right|^{2k},$$

where $h \in \mathbb{Z} \cap [1, m+1]$. If $h \geq 2$, then the local problems at the prime 2 should be resolved. However, we will see that this is not sufficient, and that the discrepancy between the integer and the polynomial analogues goes even deeper.

First, let us consider the case $h = m+1$ in order to convince the reader that resolving the constraints at the prime 2 is not sufficient. It is known that a positive proportion of polynomials $N \in \mathbb{F}_q[t]$ of degree at most r have a simple zero over \mathbb{F}_q , and that the number of zeroes of such a polynomial over \mathbb{F}_q is, on average, bounded. So we should expect

$$\text{Poly}_q(n, m, m+1; k) \asymp \frac{1}{q^n} \sum_{\substack{N \in \mathbb{F}_q[t] \\ \deg N = n}} \sum_{\substack{\alpha \in \mathbb{F}_q \\ N(\alpha)=0 \\ N'(\alpha) \neq 0}} \left| \sum_{\substack{M|N \\ \deg M \leq m}} \mu(M) \right|^{2k}.$$

If N has a simple zero at α , then we can factor $N(x) = (x - \alpha)\tilde{N}(x)$, where $\tilde{N}(\alpha) \neq 0$, that is to say $x - \alpha$ and \tilde{N} are co-prime. Then

$$\sum_{\substack{M|N \\ \deg M \leq m}} \mu(M) = \sum_{\substack{M|\tilde{N} \\ \deg M \leq m}} \mu(M) + \sum_{\substack{M|\tilde{N} \\ 1 + \deg M \leq m}} \mu((x - \alpha) \cdot M) = \sum_{\substack{M|\tilde{N} \\ \deg M = m}} \mu(M),$$

so that

$$\text{Poly}_q(n, m, m+1; k) \asymp \frac{1}{q^n} \sum_{\alpha \in \mathbb{F}_q} \sum_{\substack{\tilde{N} \in \mathbb{F}_q[t] \\ \deg \tilde{N} = n-1 \\ \tilde{N}(\alpha) \neq 0}} \left| \sum_{\substack{M|\tilde{N} \\ \deg M = m}} \mu(M) \right|^{2k} \asymp m^{2^{2k-1} - 2k-1} + 1$$

for $n \geq 2mk$, by an easy variation of Theorem 1.2. This argument can be made rigorous; we leave this task to the interested reader.

Let us now study $\text{Poly}_q(n, m, h; k)$ more generally. For any $h \in \mathbb{Z}_{\geq 1}$ and $n \geq 2mk$, we note that

$$\begin{aligned} \text{Poly}_q(n, m, h; k) &= \sum_{\substack{G_1, \dots, G_{2k} \\ m-h < \deg(G_i) \leq m \\ 1 \leq i \leq 2k}} \frac{\mu(G_1) \cdots \mu(G_{2k})}{q^{\deg([G_1, \dots, G_{2k}])}} \\ &= \sum_{\substack{G_I \ (I \in \mathcal{S}^*(2k)) \\ m-h < \sum_{I \ni i} \deg(G_I) \leq m \ (1 \leq i \leq 2k)}} \left(\prod_{I \in \mathcal{S}^-(2k)} \frac{\mu(G_I)}{q^{\deg(G_I)}} \right) \left(\prod_{I \in \mathcal{S}^+(2k)} \frac{\mu^2(G_I)}{q^{\deg(G_I)}} \right), \end{aligned}$$

as before. Applying Fourier inversion $2k$ times, we find that, for any $r \in (0, 1)$,

$$\begin{aligned} \text{Poly}_q(n, m, h; k) &= \sum_{\substack{m-h < \ell_j \leq m \\ 1 \leq j \leq 2k}} \sum_{G_I \ (I \in \mathcal{S}^*(2k))} \left(\prod_{I \in \mathcal{S}^-(2k)} \frac{\mu(G_I)}{q^{\deg(G_I)}} \right) \left(\prod_{I \in \mathcal{S}^+(2k)} \frac{\mu^2(G_I)}{q^{\deg(G_I)}} \right) \\ &\quad \times \prod_{j=1}^{2k} \int_0^1 (re(\theta_j))^{-\ell_j + \sum_{I \ni j} \deg(G_I)} d\theta_j. \end{aligned}$$

So, if we set

$$\mathcal{Z}_q(w) = \sum_{G \in \mathbb{F}_q[t]} \left(\frac{w}{q} \right)^{\deg(G)} = \prod_P (1 - (w/q)^{\deg(P)})^{-1},$$

the $\mathbb{F}_q[t]$ analogue of the Riemann zeta function, then

$$\begin{aligned} \text{Poly}_q(n, m, h; k) &= \sum_{\substack{m-h < \ell_j \leq m \\ 1 \leq j \leq 2k}} \int_{[0,1]^{2k}} \tilde{F}_q((re(\theta_j))_j) \frac{\prod_{I \in \mathcal{S}^+(2k)} \mathcal{Z}_q(r^{\#I} e(\theta_I))}{\prod_{I \in \mathcal{S}^-(2k)} \mathcal{Z}_q(r^{\#I} e(\theta_I))} \prod_{j=1}^{2k} \frac{e(-\ell_j \theta_j)}{r^{\ell_j}} d\theta \\ &= \int_{[0,1]^{2k}} \tilde{F}_q((re(\theta_j))_j) \frac{\prod_{I \in \mathcal{S}^+(2k)} \mathcal{Z}_q(r^{\#I} e(\theta_I))}{\prod_{I \in \mathcal{S}^-(2k)} \mathcal{Z}_q(r^{\#I} e(\theta_I))} \prod_{j=1}^{2k} \sum_{\ell=m-h+1}^m \frac{e(-\ell \theta_j)}{r^\ell} d\theta, \end{aligned}$$

where $\theta_I = \sum_{j \in I} \theta_j$ and $\tilde{F}_q(w)$ is a certain function that is analytic and non-zero when $|w_j| < \sqrt{q}/2k$ for all j .

We take $r = 1 - 1/m$ and note that the main contribution to $\text{Poly}_q(n, m, h; k)$ should come from those values of θ for which there are many $I \in \mathcal{S}^+(2k)$ such that $\theta_I \equiv 0 \pmod{1}$. This is the key difference with the integer case: before, we needed many $I \in \mathcal{S}^+(2k)$ with $s_I = 0$. So we see two different linear algebra problems: one over the group \mathbb{R}/\mathbb{Z} , which has torsion, and one over \mathbb{R} , which does not. The presence of torsion in \mathbb{R}/\mathbb{Z} is a reflection of the discreteness of the polynomial setting (of the degree of the polynomials, more precisely), and the fact that \mathbb{R} is a field reflects the continuous nature of the integer problem (of the logarithms of integers, more precisely).

When $h = 1$, then the integrand is $\asymp 1/m$ when $\theta_j = O(1/m) \pmod{1}$ for all j , much like the integer analogue. However, if we take $\theta_j = 1/2 + O(1/m)$ for all j , then we see that $\theta_I = O(1/m) \pmod{1}$ for $I \in \mathcal{S}^+(2k)$, whereas $\theta_I = 1/2 + O(1/m) \pmod{1}$ for $I \in \mathcal{S}^-(2k)$, so the integrand has size $m^{2^{2k-1}-1}$ for such θ . The volume of this region is $\asymp 1/m^{2k}$, leading to a contribution of size $m^{2^{2k-1}-2k-1}$ to $\text{Poly}_q(n, m, 1; k)$, which is precisely its order of magnitude for $k \geq 2$. Note that

the fact the main contribution comes from when $\theta_j \approx 1/2$ and not when $\theta_j \approx 0$ is a reflection of the local constraint at the prime 2 we noticed above.

Similarly to the above case, if $h = 2$ and $\theta_j = 1/2 + O(1/m)$, then the integrand becomes

$$\tilde{F}_q(1/2, \dots, 1/2) \frac{m^{2^{2k-1}-1}(1 + O(1/m))}{\mathcal{Z}_q(1/2)^{4^k}} \prod_{j=1}^{2k} (1 + e(\theta_j)).$$

By Taylor expansion, we have that

$$1 + e(\theta_j) = 1 - e(\theta_j - 1/2) = -(\theta_j - 1/2) - \frac{(\theta_j - 1/2)^2}{2} - \dots$$

By symmetry, we should then have that

$$\begin{aligned} & \int \cdots \int_{\substack{|\theta_j - 1/2| \leq 1/m \\ 1 \leq j \leq 2k}} \tilde{F}_q(1/2, \dots, 1/2) \frac{m^{2^{2k-1}-1}(1 + O(1/m))}{\mathcal{Z}_q(1/2)^{4^k}} \prod_{j=1}^{2k} (1 + e(\theta_j)) d\theta \\ &= (1 + O(1/m)) \int \cdots \int_{\substack{|\theta_j - 1/2| \leq 1/m \\ 1 \leq j \leq 2k}} \tilde{F}_q(1/2, \dots, 1/2) \frac{m^{2^{2k-1}-1}}{\mathcal{Z}_q(1/2)^{4^k}} \prod_{j=1}^{2k} \frac{(\theta_j - 1/2)^2}{2} d\theta, \end{aligned}$$

which leads to a contribution of size $m^{2^{2k-1}-6k-1}$ to $\text{Poly}_q(n, m, 2; k)$. This should be the dominant contribution for large k , even though for small k other regions can dominate. For example, if $k = 2$ and we take $\theta_1, \theta_2 \in [0.33, 0.34]$, $\theta_3, \theta_4 \in [0.66, 0.67]$, and $\theta_1 + \theta_2, \theta_3 + \theta_4, \theta_1 + \theta_4 = O(1/m)$, then the integrand becomes $\asymp m^5$, and we are integrating over a region of volume $\asymp 1/m^3$, so we see that $\text{Poly}_q(n, m, 1; 2) \gg m^2$. In fact, this is the exact order of magnitude of $\text{Poly}(n, m, 1; 2)$.

We conclude our discussion with another peculiar fact: if $h = 3$, then

$$\sum_{\ell=m-h+1}^m e(\ell\theta) = e(m\theta)(1 + e(\theta) + e(2\theta)) = e(m\theta) + O(1/m)$$

when $\theta = 1/2 + O(1/m)$. So $\text{Poly}_q(n, m, 3; k)$ should have the same size as $\text{Poly}_q(n, m, 1; k)$, whereas $\text{Poly}_q(n, m, h; k)$ is a bit smaller, by a factor of size $m^{O(k)}$. In general, no matter how we choose h , we cannot make the sum $\sum_{\ell=m-h+1}^m e(\ell\theta)$ small enough to cancel the contribution of the factors $\mathcal{Z}_q(r^{\#I}e(\theta_I))$ for even I in the region $\theta_j \sim 1/2$, so the quantities $\text{Poly}_q(n, m, h; k)$ do not behave in the same way as $\mathcal{M}_{\tilde{f}_0, 2k}(R)$ for k large.

3. THE ANALOGY FOR PERMUTATIONS

Completion of the proof of Theorem 1.1. It remains to prove the two claims for the quantity $c(m, k)$, which we recall is defined as the number of $(2^{2k} - 1)$ -tuples $(r_I)_{\emptyset \neq I \subset [2k]}$ of non-negative integers such that $r_I \in \{0, 1\}$ for $\#I$ odd and such that $\sum_{I: i \in I} r_I = m$, for each $i \in [2k]$.

Given any vector $\{r_I : \emptyset \neq I \subset [2k]\}$ counted by $c(m, k)$, the vector $\{r'_I : \emptyset \neq I \subset [2k]\}$ is counted by $c(m + 1, k)$ where $r'_{\{1,2\}} = r_{\{1,2\}} + 1$ and $r'_{\{3,4,\dots,2k\}} = r_{\{3,4,\dots,2k\}} + 1$, and $r'_I = r_I$ otherwise. Since $r_I \mapsto r'_I$ is injective, we see $c(m, k) \leq c(m + 1, k)$ for all $m \geq 0$, as claimed.

We now estimate $c(m, k)$. When $k = 1$, we find immediately that $c(m, 1) = 2$, so there is nothing to prove. Assume now that $k \geq 2$. Since $c(m, k)$ is increasing in m and $c(0, k) = 1$, we may assume that m is even and large enough. We note that there are $\asymp_k 1$ possibilities for the r_I

for the odd-sized I . Otherwise we have to satisfy $2k$ equations with $2^{2k-1} - 1$ variables. Hence the number of solutions should be

$$\asymp_k m^{2^{2k-1}-2k-1} + 1,$$

as claimed. Certainly, this argument yields an appropriate upper bound. To prove the lower bound for $k \geq 2$ we will construct this number of solutions. Let

$$\mathcal{I} = \{\{i, j\} : 1 \leq i < j \leq 4\} \cup \{\{1, j\} : 5 \leq j \leq 2k\},$$

so that $\#\mathcal{I} = 2k + 2$. Set $r_I = 0$ if $I \in \mathcal{S}^-(2k)$ and, given $\delta > 0$ to be chosen later, let r_I be any even integer from the range $[0, \delta m/4^k]$ if $I \in \mathcal{S}^+(2k) \setminus (\mathcal{I} \cup \{\{5, 6, \dots, 2k\}\})$. Finally, if $k > 2$, let $r_{\{5, 6, \dots, 2k\}}$ be an even integer from the range $[m - 2\delta m, m - \delta m]$. There are $\asymp_{k, \delta} m^{2^{2k-1}-3-2k}$ such choices of r_I , $I \in \mathcal{S}^+(2k) \setminus \mathcal{I}$. Then select

$$r_{\{1, j\}} := m - \sum_{I \in \mathcal{S}^+(2k) \setminus \mathcal{I}, j \in I} r_I \quad (5 \leq j \leq 2k),$$

which is an even integer lying in the interval $[0, 2\delta m]$, so that $\sum_{I \in \mathcal{S}^+(2k), j \in I} r_I = m$ for $5 \leq j \leq 2k$. Now set

$$\begin{aligned} \mathcal{I}_2 &= \{\{i, j\} : 1 \leq i < j \leq 4\}, \\ m_j &= m - \sum_{I \in \mathcal{S}^+(2k) \setminus \mathcal{I}_2, j \in I} r_I \quad (1 \leq j \leq 4). \end{aligned}$$

We note that the m_j are even integers lying in the interval $[m - 2k\delta m, m]$. It remains to choose r_I , $I \in \mathcal{I}_2$, such that $\sum_{I \in \mathcal{I}_2, j \in I} r_I = m_j$ for $1 \leq j \leq 4$. Then, we select any even integers $r_{\{2, 4\}}, r_{\{3, 4\}}$ from $[\sqrt{\delta}m - \delta m, \sqrt{\delta}m + \delta m]$, and we set

$$r_{\{1, 4\}} = m_4 - r_{\{2, 4\}} - r_{\{3, 4\}}.$$

Finally, we define $r_{\{1, 2\}}, r_{\{1, 3\}}, r_{\{2, 3\}}$ such that

$$\begin{aligned} r_{\{1, 2\}} + r_{\{1, 3\}} &= m_1 - r_{\{1, 4\}} = m_1 - m_4 + r_{\{2, 4\}} + r_{\{3, 4\}}; \\ r_{\{1, 2\}} + r_{\{2, 3\}} &= m_2 - r_{\{2, 4\}}; \text{ and} \\ r_{\{1, 3\}} + r_{\{2, 3\}} &= m_3 - r_{\{3, 4\}}. \end{aligned}$$

Note that the right-hand sides are all even so there is no parity problem, and the solutions we obtain are non-negative integers for δ small enough. We have thus constructed $\gg_k m^{2^{2k-1}-2k-1}$ solutions counted by $c(m, k)$. This completes the proof of the lemma. \square

Remark 3.1. It should not be too difficult to determine $c(m, k)$ exactly in some special cases. For example, we have that $c(m, 1) = 2$ and $c(m, 2) = \frac{1}{3}(64m^3 - 135m^2 + 182m - 66)$ for all $m \geq 1$.

Finally, we prove a probabilistic interpretation for $c(m, k)$. In its statement, we have set with a slight abuse of notation

$$(3.1) \quad M(\mathbf{c}; r) := \sum_{\substack{0 \leq b_j \leq c_j \\ 1 \leq j \leq m \\ \sum_j j b_j = r}} (-1)^{b_1 + \dots + b_m}$$

for an m -tuple of non-negative integers $\mathbf{c} = (c_1, \dots, c_m)$.

Proposition 3.1. *Let $\mathbf{X} = (X_1, X_2, \dots, X_m)$ be a vector of pairwise independent Poisson random variables, where X_j has parameter $1/j$. For every $k \in \mathbb{Z}_{\geq 1}$, we have that*

$$c(m, k) = \mathbb{E}[M(\mathbf{X}; m)^{2k}].$$

In passing, we note that Proposition 3.1 is purely a statement about Poisson random variables and not immediately related to permutations or polynomials over finite fields, but our proof makes use of this connection.

Before we prove Proposition 3.1, we need a lemma.

Lemma 3.2. *Let $N \geq m \geq 1$. The proportion of permutations $\sigma \in S_N$ that have no cycles of length $\leq m$ is*

$$\prod_{j=1}^m e^{-1/j} + O\left(\frac{m^2}{N}\right).$$

Proof. Note that this lemma was proven by the first author in [11] for large m , but here we are mainly interested in the case when m is very small compared to N . We apply inclusion-inclusion. If \mathcal{C}_j denotes the j -cycles in S_N , we write $|\pi| = j$ for an element π of \mathcal{C}_j , and we let \mathcal{C} be the union of $\mathcal{C}_1, \dots, \mathcal{C}_m$. Then

$$\begin{aligned} & \#\{\sigma \in S_N : \sigma \text{ has no cycles of length } \leq m\} \\ &= N! - \sum_{\pi \in \mathcal{C}} (N - |\pi|)! + \sum_{\substack{\pi_1, \pi_2 \in \mathcal{C} \\ \pi_1, \pi_2 \text{ disjoint}}} (N - |\pi_1| - |\pi_2|)! \mp \dots \\ &= \sum_{\substack{c_1, \dots, c_m \geq 0 \\ c_1 + 2c_2 + \dots + mc_m \leq n}} (-1)^{c_1 + \dots + c_m} (N - c_1 - 2c_2 - \dots - mc_m)! \sum_{\substack{\pi_1, \pi_2, \dots \in \mathcal{C} \text{ disjoint} \\ \#\{i: |\pi_i| = j\} = c_j \ \forall j}} 1. \end{aligned}$$

In order to count the inner quantity, we note that if $r = c_1 + \dots + c_m$ is the total number of disjoint cycles we are choosing, and we have fixed our choice for $\pi_1, \pi_2, \dots, \pi_{r-1}$, then there are

$$\frac{(N - |\pi_1| - \dots - |\pi_{r-1}|)!}{|\pi_r|! (N - |\pi_1| - \dots - |\pi_{r-1}| - |\pi_r|)!}$$

choices for the set of size $|\pi_r|$ fixed by π_r , and then $(|\pi_r| - 1)!$ possibilities for a cycle on $|\pi_r|$ given elements. Inductively, we then find that the total number of possibilities for π_1, \dots, π_r should be

$$\frac{N!}{(N - |\pi_1| - \dots - |\pi_r|)!} \cdot \frac{1}{|\pi_1| \dots |\pi_r|} = \frac{N!}{(N - c_1 - 2c_2 - \dots - mc_m)!} \prod_{j=1}^m \frac{1}{j^{c_j}}.$$

Note though we have overcounted: each possibility of j -cycles occurs $c_j!$ times, depending on the order they are picked, so we must divide the above expression by $c_1! \dots c_m!$. We then find that

$$\begin{aligned} \frac{\#\{\sigma \in S_N : \sigma \text{ has no cycles of length } \leq m\}}{N!} &= \sum_{\substack{c_1, \dots, c_m \geq 0 \\ c_1 + 2c_2 + \dots + mc_m \leq n}} \prod_{j=1}^m \frac{(-1/j)^{c_j}}{c_j!} \\ &= \prod_{j=1}^m e^{-1/j} + O\left(\frac{m^2}{N}\right), \end{aligned}$$

where the error term is obtained by noting that $\mathbf{1}_{c_1 + 2c_2 + \dots + mc_m \leq N} \leq (c_1 + 2c_2 + \dots + mc_m)/N$. \square

Proof of Proposition 3.1. We recall that we have already proved that

$$c(m, k) = \frac{1}{N!} \sum_{\sigma \in S_N} \left(\sum_{\substack{T \subset [n] \\ \sigma(T)=T \\ \#T=m}} \mu(\sigma|_T) \right)^{2k}$$

for any $N \geq 2mk$. We will now rewrite the right hand side for n much larger than m and k . Note that if σ has c_j cycles of length j for each $j \in \{1, \dots, m\}$, then

$$\sum_{\substack{T \subset [n] \\ \sigma(T)=T \\ \#T=m}} \mu(\sigma|_T) = M(\mathbf{c}; m) = \sum_{\substack{0 \leq b_j \leq c_j \\ 1 \leq j \leq m \\ \sum_j j b_j = m}} (-1)^{b_1 + \dots + b_m}$$

where $\mathbf{c} = (c_1, \dots, c_m)$. Moreover, a generalization of Cauchy's formula (see Lemma 2.2 in [5]) implies that if $t := c_1 + 2c_2 + \dots + mc_m \leq N$, then

$$\begin{aligned} & \frac{\#\{\sigma \in S_N : \sigma \text{ has } c_j \text{ } j\text{-cycles of length } j \text{ } (1 \leq j \leq m)\}}{N!} \\ &= \left(\prod_{j=1}^m \frac{1}{j^{c_j} c_j!} \right) \cdot \frac{\#\{\sigma \in S_{N-t} : \sigma \text{ has no cycles of length } \leq m\}}{(N-t)!}. \end{aligned}$$

Applying Lemma 3.2, it is then easy to conclude that

$$\begin{aligned} c(m, k) &= \lim_{N \rightarrow \infty} \frac{1}{N!} \sum_{\sigma \in S_N} \left(\sum_{\substack{T \subset [n] \\ \sigma(T)=T \\ \#T=m}} \mu(\sigma|_T) \right)^{2k} \\ &= \sum_{c_1, \dots, c_m \geq 0} M(\mathbf{c}; m)^{2k} \prod_{j=1}^m \frac{e^{-1/j}}{j^{c_j} c_j!}. \end{aligned}$$

Since $\mathbb{P}(X_j = c_j) = e^{-1/j} / (j^{c_j} c_j!)$, this is $\mathbb{E}[M(\mathbf{X}; m)^{2k}]$, and so completes the proof. \square

4. THE ANALOGY FOR POLYNOMIALS OVER FINITE FIELDS

Proof of Theorem 1.2. Throughout this proof all polynomials we consider are monic, and P denotes a generic monic irreducible polynomial over \mathbb{F}_q . Note that

$$\begin{aligned} \text{Poly}_q(n, m; k) &= \frac{1}{q^n} \sum_{\deg(F)=n} \left(\sum_{\substack{G|F \\ \deg(G)=m}} \mu(G) \right)^{2k} \\ &= \prod_{\deg(P) \leq m} \left(1 - q^{-\deg(P)} \right) \sum_{P|F \Rightarrow \deg(P) \leq m} \frac{1}{q^{\deg(F)}} \left(\sum_{\substack{G|F \\ \deg(G)=m}} \mu(G) \right)^{2k} \end{aligned}$$

for $n \geq 2km$, as can be proven by expanding the $2k$ -th power in both sides, and noticing that if $G_j|F$ for each $j \leq 2k$, then we may write $F = [G_1, \dots, G_{2k}]H$ for some monic polynomial H .

Next, note that if $F = P_1^{n_1} \cdots P_r^{n_r}$ is the factorisation of F into monic irreducible factors, and we write $c_j = \#\{i : \deg(P_i) = j\}$ for $1 \leq j \leq m$, then

$$\sum_{\substack{G|F \\ \deg(G)=m}} \mu(G) = M(\mathbf{c}; m)$$

with $M(\mathbf{c}; m)$ defined by (3.1). In particular, we see that $\sum_{G|F, \deg(G)=m} \mu(G)$ is a function of the vector $c(F) := (c_1, \dots, c_m)$. Moreover, given a fixed vector \mathbf{c} , we see that

$$\begin{aligned} \prod_{\deg(P) \leq m} (1 - q^{-\deg(P)}) \sum_{\substack{F: c(F)=\mathbf{c} \\ P|F \Rightarrow \deg(P) \leq m}} \frac{1}{q^{\deg(F)}} &= \prod_{\deg(P) \leq m} (1 - q^{-\deg(P)}) \prod_{j=1}^m \frac{\binom{N_j}{c_j}}{(q^j - 1)^{c_j}} \\ &= \prod_{j=1}^m \frac{\binom{N_j}{c_j} (1 - q^{-j})^{N_j}}{(q^j - 1)^{c_j}}, \end{aligned}$$

where

$$N_j := \#\{P \in \mathbb{F}_q[t] : P \text{ irreducible}, \deg(P) = j\}.$$

(Note that we have $(q^j - 1)^{c_j}$ and not q^{jc_j} in the denominator because we have to sum over powers of P_j too.) Galois theory implies that $q^j = \sum_{j'|j} j' N_{j'}$, whence

$$(4.1) \quad N_j = \frac{1}{j} \sum_{j'|j} \mu(j') q^{j/j'} = \frac{q^j}{j} \left(1 + O\left(\frac{\mathbf{1}_{j \geq 2}}{(q^j/j)^{1/2}}\right) \right)$$

and

$$(4.2) \quad q + j N_j \leq q^j \quad (j \geq 2).$$

Our next task is to control the quantity

$$\prod_{j=1}^m \frac{\binom{N_j}{c_j} (1 - q^{-j})^{N_j}}{(q^j - 1)^{c_j}}$$

and remove the dependence on q . First, note that

$$\prod_{j=1}^m (1 - q^{-j})^{N_j} = (1 + O(1/q)) \prod_{j=1}^m e^{-1/j}.$$

Furthermore,

$$\frac{\binom{N_j}{c_j}}{(q^j - 1)^{c_j}} = \frac{N_j^{c_j} (1 + O(c_j/N_j))^{c_j}}{c_j! (q^j - 1)^{c_j}} = \frac{1}{c_j! j^{c_j}} \left(1 + O\left(\frac{\mathbf{1}_{j \geq 2} \cdot c_j}{(q^j/j)^{1/2}} + \frac{\mathbf{1}_{j=1} c_j}{q^j}\right) \right),$$

provided that $c_1 \leq q$ and that $c_j \leq \sqrt{q^j/j}$ if $j \geq 2$. Therefore, if $c_1 \leq q$ and $\sum_{2 \leq j \leq m} c_j \cdot (j/q^j)^{1/2} \leq 1$, then

$$\prod_{j=1}^m \frac{\binom{N_j}{c_j} (1 - q^{-j})^{N_j}}{(q^j - 1)^{c_j}} = \left(1 + O\left(\frac{c_1 + 1}{q} + \sum_{j=2}^m \frac{c_j j^{1/2}}{q^{j/2}}\right) \right) \prod_{j=1}^m \frac{e^{-1/j}}{c_j! j^{c_j}}.$$

Together with Proposition 3.1, this implies that

$$\text{Poly}_q(n, m; k) = c(m, k) + O(R_1 + R_2 + R_3) \quad (n \geq 2mk),$$

where

$$\begin{aligned}
R_1 &= \sum_{c_1, \dots, c_m \geq 0} M(\mathbf{c}; m)^{2k} \left(\frac{c_1 + 1}{q} + \sum_{j=2}^m \frac{c_j j^{1/2}}{q^{j/2}} \right) \prod_{j=1}^m \frac{e^{-1/j}}{c_j! j^{c_j}}, \\
R_2 &= \sum_{\substack{c_1, \dots, c_m \geq 0 \\ c_1 > q \text{ or } \sum_{j>1} c_j j^{1/2}/q^{j/2} > 1}} M(\mathbf{c}; m)^{2k} \prod_{j=1}^m \frac{e^{-1/j}}{c_j! j^{c_j}} \\
&\leq \sum_{c_1, \dots, c_m \geq 0} M(\mathbf{c}; m)^{2k} \left(\frac{c_1}{q} + \sum_{j=2}^m \frac{c_j j^{1/2}}{q^{j/2}} \right) \prod_{j=1}^m \frac{e^{-1/j}}{c_j! j^{c_j}} \leq R_1,
\end{aligned}$$

and

$$R_3 = \sum_{\substack{c_1, \dots, c_m \geq 0 \\ c_1 > q \text{ or } \sum_{j>1} c_j j^{1/2}/q^{j/2} > 1}} M(\mathbf{c}; m)^{2k} \prod_{j=1}^m \frac{e^{-1/j} \binom{N_j}{c_j}}{(q^j - 1)^{c_j}}.$$

For R_3 , we note that $c_j \leq N_j$ in its range; otherwise, $\binom{N_j}{c_j} = 0$. In particular, $c_1 \leq N_1 = q$. Moreover, (4.2) implies that

$$\binom{N_j}{c_j} \leq \frac{N_j^{c_j}}{c_j!} \leq \begin{cases} \frac{(q^j - 1)^{c_j}}{c_j! j^{c_j}} & \text{if } j \geq 2, \\ \frac{q^{c_1}}{c_1!} \leq (1 - 1/q)^{-q} \cdot \frac{(q - 1)^{c_1}}{c_1!} & \text{if } j = 1, \end{cases}$$

Therefore

$$R_3 \ll \sum_{c_1, \dots, c_m \geq 0} M(\mathbf{c}; m)^{2k} \left(\sum_{j=2}^m \frac{c_j j^{1/2}}{q^{j/2}} \right) \prod_{j=1}^m \frac{e^{-1/j}}{c_j! j^{c_j}} \leq R_1.$$

We thus see that Theorem 1.2 is reduced to proving that $R_1 \ll_k c(m, k)/q$. It suffices to show that

$$T_i := \sum_{c_1, \dots, c_m \geq 0} c_i M(\mathbf{c}; m)^{2k} \prod_{j=1}^m \frac{e^{-1/j}}{j^{c_j} c_j!} \ll c(m, k) \quad (1 \leq i \leq m).$$

Indeed, we note that the term with $c_i = 0$ does not contribute, and we replace c_i by $c_i + 1$ to find that

$$T_i = \frac{1}{i} \sum_{c_1, \dots, c_m \geq 0} M(\mathbf{e}_i + \mathbf{c}; m)^{2k} \prod_{j=1}^m \frac{e^{-1/j}}{j^{c_j} c_j!},$$

where \mathbf{e}_i denotes the m -th dimensional vector all of whose coordinates are 0 except for the i -th coordinate that equals 1. Note that

$$M(\mathbf{e}_i + \mathbf{c}; m) = \sum_{\substack{0 \leq b_j \leq c_j \quad \forall j \neq i \\ 0 \leq b_i \leq c_i + 1 \\ \sum_j j b_j = m}} (-1)^{b_1 + \dots + b_m} = M(\mathbf{c}; m) + (-1)^{c_i + 1} M(\mathbf{c}_i; m - i(c_i + 1)),$$

where $\mathbf{c}_i = (c_1, \dots, c_{i-1}, 0, c_{i+1}, \dots, c_m)$, so that

$$M(\mathbf{e}_i + \mathbf{c}; m)^{2k} \leq 2^{2k-1} (M(\mathbf{c}; m) + M(\mathbf{c}_i; m - i(c_i + 1)))^{2k},$$

by Hölder's inequality. We thus conclude that

$$\begin{aligned} T_i &\leq \frac{2^{2k-1}}{i} c(m, k) + \frac{2^{2k-1}}{i} \sum_{c_i=0}^{\infty} \frac{e^{-1/i}}{c_i! i^{c_i}} \sum_{(c_j)_{j \leq m, j \neq i}} M(\mathbf{c}_i; m - i(c_i + 1))^{2k} \prod_{j \neq i} \frac{e^{-1/j}}{c_j! j^{c_j}} \\ &\leq \frac{2^{2k-1}}{i} c(m, k) + \frac{2^{2k-1}}{i} \sum_{c_i=0}^{\infty} \frac{e^{-1/i}}{c_i! i^{c_i}} c(m - i(c_i + 1), k), \end{aligned}$$

since the $M(\mathbf{c}_i; m - i(c_i + 1))^{2k}$ is independent of the value of the c_j 's with $j > m - i(c_i + 1)$. Recalling that $c(\ell, k)$ is an increasing function of ℓ by Theorem 1.1, we arrive to the claimed bound $T_i \ll_k c(m, k)$, whence Theorem 1.2 follows. \square

5. THE SUPPORT OF $M_{f_0}(n; R)$

We prove here (1.4), which we recall is the statement that

$$\#\{n \leq x : M_{f_0}(n; R) \neq 0\} \asymp \frac{x}{(\log R)^\delta (\log \log R)^{3/2}} \quad (x \geq R^4).$$

The lower bound was proven in the introduction, so we are left to show the upper bound. We recall the relation (1.5)

$$M_f(n; R) = \sum_{d|p_2 \cdots p_r m} \mu(d) \left\{ f\left(\frac{\log d}{\log R}\right) - f\left(\frac{\log p_1}{\log R} + \frac{\log d}{\log R}\right) \right\},$$

where $n = p_1^{\alpha_1} \cdots p_r^{\alpha_r} m$, where $p_1 < \cdots < p_r$, $\alpha_i \geq 1$ and all of the prime divisors of m are $> p_r$. Taking $r = 2$, letting q be the smallest prime dividing n and writing $n = q^j m$ with $q \nmid m$, we see that

$$M_{f_0}(n; R) = \sum_{\substack{d|m \\ R/q < d \leq R}} \mu(d).$$

Therefore,

$$(5.1) \quad \#\{n \leq x : M_{f_0}(n; R) \neq 0\} \leq \sum_{q^j \leq y} H(x/q^j, q; R/q, R) + O\left(\frac{x}{\log y}\right),$$

for any parameter $y \leq R^{1/3}$ to be chosen later, where

$$H(X, Y; Z, W) := \#\{n \leq X : P^-(n) > Y, \exists d|n \text{ with } Z < d \leq W\}.$$

We have the following estimate, that is useful in its own right.

Proposition 5.1. *Uniformly for $1 \leq Y \leq Z \leq W \leq X/(2Z)$ and $2Z \leq W \leq Z^2$, we have*

$$H(X, Y; Z, W) \ll \frac{X}{\log Y} \cdot \frac{1}{\lambda^\delta (1 + \log \lambda)^{3/2}},$$

where λ is defined by the relation $W = Z^{1+1/\lambda}$ and $\delta = 1 - \frac{1+\log \log 2}{\log 2} = 0.086071 \dots$

Remark. In the special case when $W = 2Z$, Ford [8] used a more refined argument and determined the exact order of magnitude of $H(X, Y; Z, W)$. The exact statement is a bit complicated, so we refer the interested reader to Ford's paper.

Proof. We adapt the proof of Lemma 6.1 in Ford's paper [7]. By a dyadic decomposition argument, it suffices to upper bound the difference $H(X, Y; Z, W) - H(X/2, Y; Z, W)$. Let n being counted by this difference, so that it can be written as $n = n_1 n_2$ with $n_1 \in (Z, W]$. We thus have that $n_2 \in (X/2W, X/Z]$. If $p = \min\{P^+(n_1), P^+(n_2)\} \in (Y, W]$, then we may write $n = apb$, where:

- (i) all prime factors of a are in (Y, p) ;
- (ii) all prime factors of b are $\geq p$ (and there is at least one such prime factor);
- (iii) there is a divisor $d|a$ such that $pd \in (Z, W] \cup (X/2W, X/Z]$.

If we set $\mathcal{L}(a; \sigma) := \bigcup_{d|a} [\log d - \sigma, \log d)$ and $\eta = \log(W/Z)$, the last condition can be also written as:

- (iii') either $\log(Z/p) \in \mathcal{L}(a; \eta)$, or $\log(X/(2Wp)) \in \mathcal{L}(a; \eta + \log 2)$.

Let $\eta' = \eta + \log 2$, and note that $\eta' \asymp \eta$ by our assumption that $W \geq 2Z$. Moreover, let $Z_1 = Z$ and $Z_2 = X/2W$, so that condition (iii') yields condition

- (iii'') $\log(Z_j/p) \in \mathcal{L}(a; \eta')$ for some $j \in \{1, 2\}$.

Finally, note that since there is $d|a$ with $dp > Z_j$, we must have that $p > Z_j/d \geq Z_j/a$. We thus conclude that we must have the condition

- (iv) $p > Q_j(a) := \max\{P^+(a), Z_j/a\}$.

Given a and p satisfying conditions (i), (iii'') and (iv), the number of $b \in (1, X/ap]$ such that $P^-(b) > p$ is $\ll X/(ap \log p)$. Indeed, notice that if there is one such b , then $X/ap \geq b \geq p$, so that the claimed estimate follows by a standard sieve bound, such as Theorem 4.3 of [10]. We thus conclude that

$$H(X, Y; Z, W) - H(X/2, Y; Z, W) \ll X \sum_{j=1}^2 \sum_{a \in \mathcal{P}(Y, W)} \frac{1}{a} \sum_{\substack{p > Q_j(a) \\ \log(Z_j/p) \in \mathcal{L}(a; \eta')}} \frac{1}{p \log p},$$

where $\mathcal{P}(Y, W)$ denotes the set of integers all of whose prime factors are in $(Y, W]$. As in the proof of Lemma 6.1 in [7], we have that the sum over p is $\ll L(a; \eta') / \log^2 Q_j(a)$, where $L(a; \sigma)$ denotes the Lebesgue measure of $\mathcal{L}(a; \sigma)$. We conclude that

$$\begin{aligned} H(X, Y; Z, W) - H(X/2, Y; Z, W) &\ll X \sum_{j=1}^2 \sum_{a \in \mathcal{P}(Y, W)} \frac{L(a; \eta')}{a \log^2 Q_j(a)} \\ &\ll \frac{X}{\log Y} \sum_{j=1}^2 \sum_{P^+(a') \leq Y} \sum_{a \in \mathcal{P}(Y, W)} \frac{L(a; \eta')}{aa' \log^2 Q_j(a)}. \end{aligned}$$

Since $Q_j(a) \geq Q_j(aa')$ and $L(a; \eta') \leq L(aa'; \eta')$, we have the estimate

$$H(X, Y; Z, W) - H(X/2, Y; Z, W) \ll \frac{X}{\log Y} \sum_{j=1}^2 \sum_{P^+(m) \leq W} \frac{L(m; \eta')}{m \log^2 Q_j(m)}.$$

Since $Z_2 = X/2W \geq Z = Z_1$, the contribution for $j = 2$ is bounded by the contribution from $j = 1$, and so it suffices to just consider $Z_j = Z$. In this case the contribution is $\ll 1/(\lambda^\delta (1 + \log \lambda)^{3/2})$ by Lemma 3.3, equation (3.8) and Lemma 3.7 of [7]. This completes the proof. \square

Proposition 5.1 implies that

$$H(x/q^j, q; R/q, R) \ll \frac{x}{q^j \log q} \cdot \left(\frac{\log q}{\log R} \right)^\delta \left(\log \frac{\log R}{\log q} \right)^{-3/2},$$

uniformly in $2 \leq q^j \leq y \leq R^{1/2}$ and $x \geq R^{5/2}$. Inserting this bound to (5.1), we deduce that

$$\#\{n \leq x : M_{f_0}(n; R) \neq 0\} \ll \frac{x}{(\log R)^\delta (\log \log R)^{3/2}} + \frac{x}{\log y}.$$

Selecting $y = \exp((\log R)^\delta (\log \log R)^{3/2})$ completes the proof of (1.4).

Remark 5.1. It is possible to construct integers n for which $M_{f_0}(n; R)$ is quite large. Indeed, let $y \geq 3$ and $k \in \mathbb{Z}_{\geq 1}$ be two parameters such that the interval $(y, 2^{1/k}y)$ contains at least $2k$ primes, and let $q_1 < \dots < q_{2k}$ be the smallest such primes. Then we set $n = 2q_1 \dots q_{2k}$ and $R = 2y^k$. By (1.5),

$$M_{f_0}(n; R) = \sum_{\substack{d|q_1 \dots q_{2k} \\ R/2 < d \leq R}} \mu(d).$$

The choice of R implies that the above sum runs over all divisors d of q_1, \dots, q_{2k} with precisely k prime factors, so that

$$M_{f_0}(n; R) = (-1)^k \binom{2k}{k}.$$

Optimizing the choice of k and y , and using the fact that there infinitely many y 's such that $\pi(y + \sqrt{y \log y}) - \pi(y) \gg \sqrt{y/\log y}$ (see, for example, [14, Exercice 5, p. 266]), we find that there exist arbitrarily large integers n such that $|M_f(n; R)| \gg n^{c/\log \log n}$, for any fixed $c < \frac{\log 2}{2}$ with $R \approx n^{1/2}$.

On the other hand, such extreme values of $M_{f_0}(n; R)$ are very rare, as Theorem 1.3 indicates.

6. INVERSION FORMULAS

Given $f : \mathbb{R} \rightarrow \mathbb{R}$, $R \geq 2$ and $s \in \mathbb{C}$, we set

$$\widehat{f}_R(s) = \int_0^\infty f\left(\frac{\log x}{\log R}\right) x^{s-1} dx = (\log R) \int_{-\infty}^\infty f(u) R^{su} du,$$

provided that the above integral converges. If f is Lebesgue measurable, supported in $(-\infty, 1]$ and bounded, which will always be the case for us, then \widehat{f}_R defines an analytic function for $\operatorname{Re}(s) > 0$. If, in addition, $f \in C^j(\mathbb{R})$ for some $j \geq 1$ and the derivatives $f', f'', \dots, f^{(j)}$ are all bounded, then integrating by parts j times yields the formula

$$(6.1) \quad \widehat{f}_R(s) = \frac{(-1)^j}{s^j (\log R)^{j-1}} \int_{-\infty}^\infty f^{(j)}(u) R^{su} du.$$

In particular, we see that

$$(6.2) \quad \left| \widehat{f}_R(s) \right| \leq \|f^{(j)}\|_\infty \cdot \frac{R^{\operatorname{Re}(s)}}{\operatorname{Re}(s)(|s| \log R)^j}$$

for $\operatorname{Re}(s) > 0$, where we used our assumption that $\operatorname{supp}(f) \subset (-\infty, 1]$.

Now, for $m \in \mathbb{Z}_{\geq 1}$, the Mellin inversion formula implies that for $c > 0$

$$(6.3) \quad f\left(\frac{\log m}{\log R}\right) = \frac{1}{2\pi i} \int_{\operatorname{Re}(s)=c} \widehat{f}_R(s) m^{-s} ds.$$

In the proof of Theorem 1.3 with $A \geq 1$ and of Theorem 1.6, our assumption that f is a few times differentiable in \mathbb{R} allows us to apply (6.2) and write $M_f(a; R)$ in terms of an absolutely convergent integral, which can easily be truncated at some appropriate height. However, when $A = 0$ in

Theorem 1.3, we have $f_0 = \chi_{(-\infty, 1]}$, so that $\widehat{(f_0)}_R(s) = R^s/s$. Truncating Perron's formula is still feasible but rather technical. Instead, we perform a technical manoeuvre and smoothen f_0 a bit. We consider a smooth function $h : \mathbb{R} \rightarrow \mathbb{R}$ such that

$$\begin{cases} h(x) = 1 & \text{if } x \leq 1 - \eta, \\ 0 \leq h(x) \leq 1 & \text{if } 1 - \eta \leq x \leq 1, \\ h(x) = 0 & \text{if } x \geq 1, \end{cases}$$

where $\eta = 1/(\log R)^C$ for some constant $C > 0$ that will be chosen appropriately later. We choose h so that $h^{(j)}(x) \ll_j \eta^{-j}$, for all $j \in \mathbb{Z}_{\geq 0}$. We claim that, for any fixed $L > 0$ and $k \geq 1$, there is $C = C(k, L)$ such that

$$(6.4) \quad \mathcal{M}_{f_0, 2k}(R) = \mathcal{M}_{h, 2k}(R) + O\left(\frac{1}{(\log R)^L}\right).$$

Indeed, we have that

$$|\mathcal{M}_{f_0, 2k}(R) - \mathcal{M}_{h, 2k}(R)| \leq 2k \sum_{\substack{d_1, \dots, d_{2k-1} \leq R \\ R^{1-\eta} < d_{2k} \leq R}} \frac{\prod_{j=1}^{2k} \mu^2(d_j)}{[d_1, \dots, d_{2k}]} \leq 2k \sum_{m \leq R^{2k}} \frac{\tau(m)^{2k-1}}{m} \sum_{\substack{d|m \\ R^{1-\eta} < d \leq R}} 1,$$

by setting $m = [d_1, \dots, d_{2k}]$ and $d = d_{2k}$. We split the above sum according to whether $\tau(m) \leq (\log R)^B$ or not, where B is some parameter. We then find that

$$\begin{aligned} |\mathcal{M}_{f_0, 2k}(R) - \mathcal{M}_{h, 2k}(R)| &\leq 2k \sum_{\substack{m \leq R^{2k} \\ \tau(m) \leq (\log R)^B}} \frac{(\log R)^{(2k-1)B}}{m} \sum_{\substack{d|m \\ R^{1-\eta} < d \leq R}} 1 \\ &\quad + 2k \sum_{\substack{m \leq R^{2k} \\ \tau(m) > (\log R)^B}} \frac{\tau(m)^{2k+1} (\log R)^{-B}}{m} \\ &\ll_k (\log R)^{(2k-1)B+2-C} + (\log R)^{2^{2k+1}-B}. \end{aligned}$$

We choose $B = L + 2^{2k+1}$ and $C \geq (2k-1)2^{2k+1} + 2kL + 2$ to complete the proof of our claim. For the purposes of Theorem 1.3, we may take $L = 1$, so that having $C \geq (2k-1)2^{2k+1} + 2k + 2$ suffices. We also note that

$$\widehat{h}_R(s) = -\frac{1}{s} \int_{1-\eta}^1 h'(u) R^{su} du = -\frac{R^s}{s} \int_0^\eta h'(1-u) R^{-su} du,$$

by (6.1) and the fact that h is constant outside $[1-\eta, 1]$. In particular, this relation implies that \widehat{h}_R has a meromorphic continuation to \mathbb{C} with only a simple pole at $s = 0$ of residue $-\int_0^\eta h'(1-u) du = 1$. We further note that

$$(6.5) \quad \begin{aligned} \frac{d^j}{ds^j} \left(\frac{s \widehat{h}_R(s)}{R^s} \right) &= (-1)^{j-1} \int_0^\eta h'(1-u) (u \log R)^j R^{-su} du \\ &\ll \eta \cdot \eta^{-1} \cdot (\eta \log R)^j = (\eta \log R)^j \quad (s \in \mathbb{C}, \operatorname{Re}(s) \geq -1). \end{aligned}$$

Moreover, we have that

$$(6.6) \quad \begin{aligned} \widehat{h}_R(s) &= (\log R) \int_{-\infty}^1 R^{us} du + O(\eta(\log R) R^{\operatorname{Re}(s)}) \\ &= \frac{R^s}{s} + O((\log R)^{-C+1} R^{\operatorname{Re}(s)}) \quad (\operatorname{Re}(s) \geq 0), \end{aligned}$$

a relation that we will use at the very end of the proof of Theorem 1.3.

7. A COMBINATORIAL PROBLEM IN LINEAR ALGEBRA

Recall the notations from Section 1.7. Consider the $2k$ -dimensional vector space (over \mathbb{Q}) of linear forms in the free variables s_1, \dots, s_{2k} , which we denote by W_k . Given a subspace V of W_k , we define

$$\mathcal{A}(V) = \sum_{\substack{J \in \mathcal{S}^*(2k) \\ s_J \in V}} (-1)^{\#J}.$$

(We recall that in our notation $s_J = \sum_{j \in J} s_j$.) We will prove the following result.

Proposition 7.1. *Let $k \geq 1$ and V be a subspace of W_k containing the form $s_{[2k]} = \sum_{i=1}^{2k} s_i$.*

- (a) *If $s_j \in V$ for some $j \in [2k]$, then $\mathcal{A}(V) = -1$.*
- (b) *If $\dim(V) = 2k - 1$, then*

$$\mathcal{A}(V) - \dim(V) \leq \binom{2k}{k} - 2k,$$

with equality if, and only if, there is a set $J \subset [2k]$ such that $\#J = k$, $1 \in J$, and $V = \operatorname{Span}_{\mathbb{Q}}(\{s_j - s_1\}_{j \in J}, \{s_j + s_1\}_{j \in [2k] \setminus J})$.

- (c) *If $\dim(V) \leq 2k - 2$, then*

$$\mathcal{A}(V) - \dim(V) \leq \binom{2k}{k} - 2k - 2.$$

Proof. (a) If $s_j \in V$, then we immediately see that $\mathcal{A}(V) = -1$ by pairing s_J with $s_{J \cup \{j\}}$ for each $J \subset [2k] \setminus \{j\}$.

(b) We may assume that $s_1, \dots, s_{2k} \notin V$, by part (a). Since $\dim(V) = 2k - 1$ and $s_1 \notin V$, for each $j = 1, \dots, 2k$ we have that $s_j \equiv r_j s_1 \pmod{V}$, for some $r_j \in \mathbb{Q} \setminus \{0\}$. We may write $r_j = b_j/q$ for some $b_j \in \mathbb{Z} \setminus \{0\}$ and $q \in \mathbb{Z}_{\geq 1}$, so that $s_J \in V$ if, and only if, $b_J = \sum_{j \in J} b_j = 0$. Therefore

$$\mathcal{A}(V) = -1 + \int_0^1 \prod_{j=1}^{2k} (1 - e(b_j \theta)) d\theta \leq -1 + \int_0^1 \prod_{j=1}^{2k} |1 - e(b_j \theta)| d\theta.$$

Hölder's inequality then implies that

$$\mathcal{A}(V) \leq -1 + \prod_{j=1}^{2k} \left(\int_0^1 |1 - e(b_j \theta)|^{2k} d\theta \right)^{\frac{1}{2k}} = -1 + \binom{2k}{k},$$

whence

$$(7.1) \quad \mathcal{A}(V) \leq \binom{2k}{k} - 1.$$

Finally, we claim that (7.1) is an equality if, and only if, the multiset $\{b_1, \dots, b_{2k}\}$ is of the form $\{b, -b, \dots, b, -b\}$ with $b = b_1$ (which must equal q). This claim immediately implies (b) of the Proposition.

If the multiset $\{b_1, \dots, b_{2k}\}$ is of the form $\{b, -b, \dots, b, -b\}$, then the integral formula for $\mathcal{A}(V)$ becomes $\mathcal{A}(V) = -1 + \int_0^1 |1 - e(b\theta)|^{2k} d\theta = \binom{2k}{k} - 1$. Conversely, we know that Hölder's inequality above is an equality if, and only if, there exist real numbers $\lambda_1, \dots, \lambda_{2k}$ such that $|1 - e(b_j\theta)| = \lambda_j |1 - e(b_1\theta)|$ for $\theta \in [0, 1]$ and $j \in \{1, \dots, 2k\}$. Since $\int_0^1 |1 - e(b\theta)| d\theta = 4/\pi$ for $b \neq 0$, we must have that $\lambda_j = 1$ for all j . Moreover, taking θ close enough to 0, we find that the condition $|1 - e(b_j\theta)| = |1 - e(b_1\theta)|$ implies that $|b_j| = |b_1|$ for all j . So $\{b_1, \dots, b_{2k}\}$ has ℓ copies of b_1 and $2k - \ell$ copies of $-b_1$, for some $\ell \in \{1, \dots, 2k\}$. Since $b_{[2k]} = 0$ by our assumption that $s_{[2k]} \in V$, we must have that $\ell = k$, which completes the proof of our claim.

(c) Write $\dim(V) = 2k - n$, where $n \geq 2$. By part (a), we may assume that $s_1, \dots, s_{2k} \notin V$.

We first deal with the case $n = 2, k = 2$ by direct computation. In this case, we have $s_1 + s_2 + s_3 + s_4 \in V$ and $s_1, \dots, s_4 \notin V$, by assumption. It is thus easy to see that either $V \cap \{s_I : I \in \mathcal{S}^*(2k)\} = \{s_1 + \dots + s_4\}$ or $V \cap \{s_I : I \in \mathcal{S}^*(2k)\} = \{s_1 + \dots + s_4, s_J\}$, for some J containing two elements. (Here we recall that $\mathcal{S}^*(2k) = \{I \subseteq [2k] : I \neq \emptyset\}$.) In any case, $\mathcal{A}(V) \leq 2$, as required. This completes the proof of part (c) when $n = 2$ and $k = 2$.

We now assume that either $n > 2$ or $k > 2$. Choose a maximal subset of linear forms $\{s_{j_1}, \dots, s_{j_{n'}}\}$ that are linearly independent when reduced mod V . Clearly, $n' = n$. Moreover, a permutation of the variables s_1, \dots, s_{2k} allows to assume without loss of generality that $j_i = i$ for each i . Then

$$s_j \equiv \sum_{i=1}^n r_{i,j} s_i \pmod{V} \quad (1 \leq j \leq 2k),$$

for certain $r_{i,j} \in \mathbb{Q}$. We write $r_{i,j} = b_{i,j}/q$, where $b_{i,j} \in \mathbb{Z}$ and $q \in \mathbb{Z}_{\geq 1}$, so that $s_J \in V$ if, and only if, $b_{i,J} := \sum_{j \in J} b_{i,j} = 0$ for each $i \in [n]$. Thus

$$\mathcal{A}(V) + 1 = \int_{[0,1]^n} \prod_{j=1}^{2k} (1 - e(b_{1,j}\theta_1 + \dots + b_{n,j}\theta_n)) d\theta_1 \dots d\theta_n.$$

We set

$$J_m = \{1 \leq j \leq 2k : b_{n-m+1,j} = \dots = b_{n,j} = 0\} \quad (0 \leq m \leq n)$$

to be the set of j such that s_j is in the span of $\{s_1, \dots, s_{n-m}\} \pmod{V}$. In particular, $J_0 = [2k]$ and $J_n = \emptyset$. By construction, s_i is a basis vector of W_k/V for $1 \leq i \leq n$, so for $0 \leq m \leq n-1$ we have $n-m \in J_m$ but $n-m \notin J_{m+1}$. In particular, $\#(J_m \setminus J_{m+1}) \geq 1$ for $0 \leq m \leq n-1$. Then

$$\begin{aligned} \mathcal{A}(V) + 1 &\leq \int_{[0,1]^{n-1}} \prod_{j \in J_1} |1 - e(b_{1,j}\theta_1 + \dots + b_{n-1,j}\theta_{n-1})| \\ &\quad \times \left(\int_0^1 \prod_{j \in [2k] \setminus J_1} |1 - e(b_{1,j}\theta_1 + \dots + b_{n,j}\theta_n)| d\theta_n \right) d\theta_1 \dots d\theta_{n-1}. \end{aligned}$$

By Hölder's inequality, the innermost integral is bounded by

$$\prod_{j \in [2k] \setminus J_1} \left(\int_0^1 |1 - e(b_{1,j}\theta_1 + \dots + b_{n,j}\theta_n)|^{2k - \#J_1} d\theta_n \right)^{\frac{1}{2k - \#J_1}}.$$

Since $b_{n,j} \neq 0$ for $j \notin J_1$, we make the change of variables $\theta_n \rightarrow b_{1,j}\theta_1 + \cdots + b_{n,j}\theta_n$ and use periodicity to find that

$$\begin{aligned} \int_0^1 |1 - e(b_{1,j}\theta_1 + \cdots + b_{n,j}\theta_n)|^{2k-\#J_1} d\theta_n &= \frac{1}{|b_{n,j}|} \int_0^{|b_{n,j}|} |1 - e(\theta)|^{2k-\#J_1} d\theta \\ &= \int_0^1 |1 - e(\theta)|^{2k-\#J_1} d\theta. \end{aligned}$$

We set

$$M(\lambda) = \int_0^1 |1 - e(\theta)|^\lambda d\theta = 2^\lambda \int_0^1 |\sin(\pi\theta)|^\lambda d\theta.$$

Note that $M(2k) = \binom{2k}{k}$. Thus we find that

$$\mathcal{A}(V) + 1 \leq M(2k - \#J_1) \int_{[0,1]^{n-1}} \prod_{j \in J_1} |1 - e(b_{1,j}\theta_1 + \cdots + b_{n-1,j}\theta_{n-1})| d\theta_1 \cdots d\theta_{n-1}.$$

We repeat the same process to obtain

$$\begin{aligned} \mathcal{A}(V) + 1 &\leq M(2k - \#J_1) M(\#J_1 - \#J_2) \\ &\quad \times \int_{[0,1]^{n-1}} \prod_{j \in J_2} |1 - e(b_{1,j}\theta_1 + \cdots + b_{n-1,j}\theta_{n-1})| d\theta_1 \cdots d\theta_{n-2} \\ &\leq M(2k - \#J_1) M(\#J_1 - \#J_2) \cdots M(\#J_{n-2} - \#J_{n-1}) M(\#J_{n-1}). \end{aligned}$$

Thus

$$(7.2) \quad \mathcal{A}(V) + 1 \leq \sup\{M(\lambda_1) \cdots M(\lambda_n) : \lambda_1 + \cdots + \lambda_n = 2k, \lambda_j \geq 1 \ (1 \leq j \leq n)\}.$$

By Cauchy-Schwarz, for any positive reals x, y we have

$$2(xy)^{(A+B)/2} = 2(xy)^B (xy)^{(A-B)/2} \leq (xy)^B (x^{A-B} + y^{A-B}) = x^A y^B + x^B y^A.$$

Thus, applying this with $x = |\sin \theta_1|$, $y = |\sin \theta_2|$ we find

$$\begin{aligned} M(\lambda_1) M(\lambda_2) &= \frac{2^{\lambda_1+\lambda_2}}{2} \int_0^1 \int_0^1 \left(|\sin(\pi\theta_1)^{\lambda_1} \sin(\pi\theta_2)^{\lambda_2}| + |\sin(\pi\theta_1)^{\lambda_1} \sin(\pi\theta_2)^{\lambda_2}| \right) d\theta_1 d\theta_2 \\ &\geq 2^{\lambda_1+\lambda_2} \left(\int_0^1 |\sin(\pi\theta)|^{(\lambda_1+\lambda_2)/2} d\theta \right)^2 = M\left(\frac{\lambda_1 + \lambda_2}{2}\right)^2. \end{aligned}$$

In particular, $\log M(\lambda)$ is a convex function. It is then easy to see that supremum in (7.2) is attained when $\lambda_j = 1$ for $n-1$ of the indices $j \in [n]$, and with the remaining λ_j being equal to $2k - n + 1$. Indeed, without loss of generality $\lambda_1, \dots, \lambda_{n-1} \leq \lambda_n$, and if $\lambda_j \neq 1$ for some $j < n$, then we can increase the size of $M(\lambda_1) \cdots M(\lambda_n)$ by replacing λ_j with $\lambda_j - 1$ and λ_n with $\lambda_n + 1$. So

$$\mathcal{A}(V) \leq M(1)^{n-1} M(2k - n + 1) - 1.$$

Thus, it suffices to show that

$$(7.3) \quad M(1)^{n-1} M(2k - n + 1) < \binom{2k}{k} - n = M(2k) - n$$

for $2 \leq n \leq 2k - 1$ and $k \geq 2$.

Firstly, consider $n = 2$ and $k \geq 3$. The function $k \mapsto M(1)M(2k-1)/M(2k)$ is decreasing in k by the convexity of $\log M(\lambda)$. Thus

$$M(1)M(2k-1) \leq \frac{M(1)M(3)}{M(4)}M(2k) = \frac{64}{9\pi^2} \binom{2k}{k} < \binom{2k}{k} - 2.$$

Here we have used the fact $M(1) = 4/\pi$, $M(3) = 32/3\pi$ and performed a quick computation to verify $64\binom{2k}{k}/(9\pi^2) < \binom{2k}{k} - 2$ for all $k \geq 3$.

Now consider $3 \leq n \leq 2k-1$. The function $n \mapsto M(1)^{n-1}M(2k-n+1)$ is decreasing in n since

$$\frac{M(1)^{n-1}M(2k-n+1)}{M(1)^{n-2}M(2k-n+2)} = \frac{M(1)M(2k-n+1)}{M(2k-n+2)} \leq \frac{M(1)^2}{M(2)} < 1.$$

Similarly $k \mapsto M(2k-2)/M(2k)$ is decreasing in k respectively by the convexity of $\log M(\lambda)$. Thus we have

$$\begin{aligned} M(1)^{n-1}M(2k-n+1) &\leq M(1)^2M(2k-2) \\ &\leq \frac{M(1)^2M(2)}{M(4)}M(2k) \\ &= \frac{16}{3\pi^2} \binom{2k}{k} \\ &< \binom{2k}{k} - 2k + 1 = M(2k) - 2k + 1. \end{aligned}$$

Here we have performed a short computation to verify the final inequality. This completes the proof of the proposition. \square

8. CONTOUR INTEGRATION

In this section we begin our attack on Theorem 1.3. All implied constants might depend on k and on A . We will actually prove a result that is a little weaker than Theorem 1.3: we will show that there exist constants $c_{k,A}$ and c'_k for which

$$(8.1) \quad \mathcal{M}_{f_A, 2k}(R) = c_{k,A}(\log R)^{\mathcal{E}_{k,A}} + O((\log R)^{\mathcal{E}_{k,A}-1}),$$

and

$$(8.2) \quad \mathcal{M}_{\tilde{f}_0, 2k}(R) = c'_k(\log R)^{\binom{2k}{k}-2k} + O((\log R)^{\binom{2k}{k}-2k-1}).$$

Here we recall that $\mathcal{E}_{k,A} = \max(\binom{2k}{k} - 2k(A+1), -1)$, and we have defined $\tilde{f}_0(x) = f_0(x) - f_0(x + \frac{\log 2}{\log R})$, so that

$$M_{\tilde{f}_0}(n; R) = \sum_{\substack{d|n \\ R/2 < d \leq R}} \mu(d).$$

Notice that we do not claim here that $c_{k,A} \neq 0$ and $c_k \neq 0$, as is required in order to prove Theorem 1.3. We do obtain a very complicated expression for these constants, but we are unable to prove they are non-zero (or evaluate them at all). Showing that $c_{k,A} > 0$ and $c'_k > 0$ is the objective of Section 9.

8.1. Initial preparations. We will first prove relation (8.1). The proof of relation (8.2) is very similar, and we indicate the necessary changes in the end of Section 8.

We note that

$$(8.3) \quad \widehat{(f_A)}_R(s) = \frac{A!R^s}{(\log R)^{A s^{A+1}}}.$$

This function is absolutely integrable on vertical lines $\operatorname{Re}(s) = c \neq 0$ when $A \geq 1$, but this is not the case when $A = 0$. However, recall from relation (6.4) that

$$\mathcal{M}_{f_0,2k}(R) = \mathcal{M}_{h,2k}(R) + O\left(\frac{1}{(\log R)^2}\right),$$

where h is a smooth function such that $h(x) = 1$ for $x \leq 1 - 1/(\log R)^C$ and $h(x) = 0$ for $x \geq 1$, for some constant $C \geq (2k-1)2^{2k+1} + 2k + 2$ to be chosen later. Therefore relation (8.1) is reduced to showing that

$$(8.4) \quad \mathcal{M}_{g,2k}(R) = c_{k,A}(\log R)^{\mathcal{E}_{k,A}} + O((\log R)^{\mathcal{E}_{k,A}-1}),$$

where $g = h$ when $A = 0$, and $g = f_A$ when $A \geq 1$.

For any $\lambda > 1$, which will be chosen to be sufficiently large in terms of k , relation (6.3) implies that

$$\mathcal{M}_{g,2k}(R) = \sum_{\substack{m_j \in \mathbb{Z}_{\geq 1} \\ 1 \leq j \leq 2k}} \frac{\prod_{j=1}^{2k} \mu(m_j)}{[m_1, \dots, m_{2k}]} \cdot \frac{1}{(2i\pi)^{2k}} \int \cdots \int_{\substack{\operatorname{Re}(s_j) = \lambda^j / \log R \\ 1 \leq j \leq 2k}} \prod_{j=1}^{2k} m_j^{-s_j} \left(\prod_{j=1}^{2k} \widehat{g}_R(s_j) \right) ds_{2k} \cdots ds_1.$$

To this end, we introduce the multiple Dirichlet series

$$D(\mathbf{s}) := \sum_{\substack{m_j \in \mathbb{Z}_{\geq 1} \\ 1 \leq j \leq 2k}} \frac{\prod_{j=1}^{2k} m_j^{-s_j} \mu(m_j)}{[m_1, \dots, m_{2k}]},$$

which converges absolutely when $\operatorname{Re}(s_j) > 0$ for all j as can be seen, for example, by the Euler product expansion

$$(8.5) \quad \begin{aligned} D(\mathbf{s}) &= \prod_p \left(\sum_{\nu_1, \dots, \nu_k \in \{0,1\}} \frac{(-1)^{\nu_1 + \dots + \nu_k}}{p^{\nu_1 s_1 + \dots + \nu_k s_k}} \cdot \frac{1}{[p^{\nu_1}, \dots, p^{\nu_k}]} \right) \\ &= \prod_p \left(1 + \frac{1}{p} \sum_{\emptyset \neq I \subset [2k]} \frac{(-1)^{\#I}}{p^{s_I}} \right) \end{aligned}$$

$$(8.6) \quad = \prod_p \left(1 - \frac{1}{p} + \frac{1}{p} \prod_{j=1}^{2k} \left(1 - \frac{1}{p^{s_j}} \right) \right),$$

where we have used the notation $s_I = \sum_{i \in I} s_i$. (Similar computations are performed in [1].) We thus see that

$$\mathcal{M}_{g,2k}(R) = \frac{1}{(2i\pi)^{2k}} \int \cdots \int_{\substack{\operatorname{Re}(s_j) = \lambda^j / \log R \\ 1 \leq j \leq 2k}} D(\mathbf{s}) \left(\prod_{j=1}^{2k} \widehat{g}_R(s_j) \right) ds_{2k} \cdots ds_1.$$

for any $\lambda > 1$.

We shall truncate all variables of integration at height

$$T := \exp\{(\log \log R)^2\}.$$

To do so, we notice that $\widehat{g}_R(s) \ll (\log R)^{O(1)}/|s|^2$ for $\operatorname{Re}(s) = \lambda^j/\log R$, a consequence of (6.2) when $A = 0$ and of (8.3) when $A \geq 1$, as well as that $D(\mathbf{s}) \ll (\log R)^{O(1)}$, an estimate that follows by formula (8.5) and the Prime Number Theorem. We conclude that

$$\mathcal{M}_{g,2k}(R) = I_{g,2k}(R) + O\left(\frac{1}{(\log R)^2}\right),$$

where

$$I_{g,2k}(R) := \frac{1}{(2i\pi)^{2k}} \int \cdots \int_{\substack{\operatorname{Re}(s_j) = \lambda^j/\log R \\ |\operatorname{Im}(s_j)| \leq T \\ 1 \leq j \leq 2k}} D(\mathbf{s}) \left(\prod_{j=1}^{2k} \widehat{g}_R(s_j) \right) ds_{2k} \cdots ds_1.$$

Motivated by (8.5) and (8.6), we set

$$\begin{aligned} P(\mathbf{s}) &:= D(\mathbf{s}) \prod_{I \in \mathcal{S}^*(2k)} \zeta(1 + s_I)^{(-1)^{1+\#I}} \\ &= \prod_p \left\{ \left(1 - \frac{1}{p} + \frac{1}{p} \prod_{j=1}^{2k} \left(1 - \frac{1}{p^{s_j}} \right) \right) \prod_{I \in \mathcal{S}^*(2k)} \left(1 - \frac{1}{p^{1+s_I}} \right)^{(-1)^{\#I}} \right\}, \end{aligned}$$

which is analytic when $\operatorname{Re}(s_j) > -1/(4k)$ for all j , as well as

$$F(\mathbf{s}) := P(\mathbf{s}) \prod_{j=1}^{2k} \frac{(\log R)^A \widehat{g}_R(s_j)}{R^{s_j} \zeta(1 + s_j)^{A+1}}$$

and

$$e_I := \begin{cases} A & \text{if } \#I = 1, \\ (-1)^{\#I} & \text{if } \#I \geq 2, \end{cases}$$

so that

$$I_{g,2k}(R) = \frac{1}{(2i\pi)^{2k}} \int \cdots \int_{\substack{\operatorname{Re}(s_j) = \lambda^j/\log R \\ |\operatorname{Im}(s_j)| \leq T \\ 1 \leq j \leq 2k}} \frac{F(\mathbf{s}) R^{s_1 + \cdots + s_{2k}}}{(\log R)^{2kA}} \prod_{I \in \mathcal{S}^*(2k)} \zeta(1 + s_I)^{e_I} ds_{2k} \cdots ds_1.$$

Given $\ell \in \mathbb{N}$, we now let

$$\Omega_\ell := \{\mathbf{s} \in \mathbb{C}^\ell : |\operatorname{Re}(s_j)| < 2/(\log T)^{4/3}, |\operatorname{Im}(s_j)| < T + 1 \ (1 \leq j \leq \ell)\}$$

and define \mathcal{C}_ℓ to be the class of complex-valued functions f such that: (a) f is defined over a complex domain containing Ω_ℓ ; (b) f is analytic in Ω_ℓ ; (c) the derivatives of f satisfy the bound

$$(8.7) \quad \frac{\partial^{j_1 + \cdots + j_\ell} f}{\partial s_1^{j_1} \cdots \partial s_\ell^{j_\ell}}(\mathbf{s}) \ll_{j_1, \dots, j_\ell} \frac{(\log \log R)^{O(j_1 + \cdots + j_\ell)}}{(|s_1| + 1) \cdots (|s_\ell| + 1)}$$

for all $j_1, \dots, j_\ell \geq 0$ and all $\mathbf{s} = (s_1, \dots, s_\ell) \in \Omega_\ell$.

We claim that $F \in \mathcal{C}_{2k}$. Indeed, there are absolute constants $\delta, c_0 > 0$ such that $\zeta(s)(s-1)$ is analytic and non-vanishing for $|s-1| \leq \delta$ and

$$(8.8) \quad \zeta^{(j)}(s), \left(\frac{1}{\zeta}\right)^{(j)}(s) \ll_j \log^{j+1}(|t|+2) \quad \text{when } \sigma \geq 1 - \frac{c_0}{\log(|t|+2)}, |s-1| \geq \delta,$$

with (8.8) being a consequence⁴ of the classical zero-free region for ζ . Moreover,

$$(8.9) \quad \frac{d^j}{ds^j} \left(\frac{s^{A+1} \widehat{g}_R(s)}{R^s} \right) \ll \frac{1}{(\log R)^A} \quad (\operatorname{Re}(s) \geq -1, j \in \mathbb{Z}_{\geq 0}),$$

an estimate that follows from (6.5) when $A = 0$ and from the formula (8.3) for $(\widehat{f_A})_R$ otherwise. Our claim that $F \in \mathcal{C}_{2k}$ then follows.

8.2. Contour shifting. We will simplify $I_{g,2k}(R)$ and prove (8.1) by a $2k$ -dimensional contour shifting argument that we will demonstrate in an iterative fashion. The general idea is to move the variables s_j to the left in a certain order. When we move the contour corresponding to the variable s_j , we will pick up contributions from poles of the integrand (coming from solutions to linear equations of the form $s_I = 0$, $I \in \mathcal{S}^*(2k)$ with $e_I > 0$), and be left with a residual contour (which will be negligible in size). Thus we only need to consider the contributions from the poles, and these contributions will all be multi-integrals similar to $I_{g,2k}(R)$ but involving one fewer variable. By iterating this, we show that $I_{g,2k}(R)$ is (up to a small error term) given by the total contribution of all the successive poles we have encountered having shifted all $2k$ variables. We will show that provided one moves the contours in a suitable order, all the contributions from all the multi-poles and all the residual integrals give a contribution $c_{k,A}(\log R)^{\varepsilon_{k,A}} + O((\log R)^{\varepsilon_{k,A}-1})$.

When we consider poles we encounter equations of the form $s_I = 0$, where we think of $s_I = \sum_{i \in I} s_i$ as a linear form in the variables s_1, \dots, s_{2k} . To avoid any ambiguity when we consider multiple such equations, we will let $L_{0,I} \in \mathbb{Q}[x_1, \dots, x_{2k}]$ be the linear form corresponding to s_I , that is to say

$$L_{0,I}(x) := \sum_{i \in I} x_i.$$

Before we setup the necessary notation to keep track of all the terms we encounter when performing the multiple contour shifting, we first describe the first two contour shifting steps to help motivate the basic idea.

The first variable we move to the left is s_{2k} . When doing so, we pick up the contribution from some poles in the integrand. Such a pole must occur when $L_{0,I_1}(s) = 0$ for some $I_1 \subset [2k]$ with $e_{I_1} > 0$ and $2k \in I_1$ (a possible pole from $\prod_{I \in \mathcal{S}^*(2k)} \zeta^{e_I}(1 + L_{0,I}(s))$.) Having fixed such a pole and the corresponding set I_1 , we use this equation $L_{0,I_1}(s) = 0$ to rewrite s_{2k} in terms of s_j , $j \in [2k] \setminus \{2k\}$. Imposing the same condition on the x_j 's, we find that for each $I \subset [2k]$, the linear form $L_{0,I}(x_1, \dots, x_{2k})$ becomes a linear form $L_{1,I}$ in the variables x_j for $j \in [2k] \setminus \{2k\}$. Trivially, $L_{1,I} = 0$ if, and only if, $I \in \mathcal{I}_1 := \{\emptyset, I_1\}$. This pole contribution can be written as an integral over s_1, \dots, s_{2k-1} , with an integrand that has poles only when $L_{1,I}(s) = 0$.

Next, for this integral over s_1, \dots, s_{2k-1} , we choose some other variable s_{j_2} (precisely how we choose s_{j_2} will be specified later), and move the s_{j_2} contour. This produces a residual contour (which will be negligible) and contributions from further poles in the integrand which occur only when s_{j_2} satisfies a linear equation $L_{1,I_2}(s) = 0$ for some $I_2 \in \mathcal{S}^*(2k) \setminus \mathcal{I}_1$ with $e_{I_2} > 0$ and with

⁴The claimed bound follows by [26, Theorems 3.8 and 3.11] and the fact that if f is analytic in a neighbourhood of the circle $|z| \leq r$, then $f^{(j)}(z_0) = \frac{1}{2\pi i} \oint_{|z|=r} f(z) dz / z$ for any z_0 with $|z_0| < r$.

$L_{1,I_2}(\mathbf{x})$ having a non-zero x_{j_2} coefficient. We use this to write s_{j_2} in terms of s_j , $j \in [2k] \setminus \{2k, j_2\}$. Imposing the corresponding condition on the variables x_j makes $L_{1,I}$ a linear form $L_{2,I}$ in the variables x_j , $j \in [2k] \setminus \{2k, j_2\}$. Some of these new linear forms will vanish identically, and the total number will determine the order of this pole.

Continuing in this manner, we eventually write our original integral $I_{g,2k}(R)$ in terms of $O(1)$ contributions from repeatedly encountered poles (all of which will be of the form $c(\log R)^m$ for some c, m) or from terms which correspond to encountering a residual integral (which will always be small). In order to control this process, we need to keep track of which poles we encounter, the order of the poles, and the integrands of the new multi-integrals corresponding to these poles. To do this we introduce some notation and terminology.

- Let us be given an integer $N \in \{0, 1, \dots, 2k\}$, which we shall often refer to as the **level**. It describes how many iterations we have performed (i.e. how many variables s_j we have shifted). The case $N = 0$ corresponds to the initial integral $I_{g,2k}(R)$.
- Let us be given sets $I_1, \dots, I_N \subset [2k]$ and indices j_1, \dots, j_N such that:
 - (i) $j_n \in I_n$ for each $1 \leq n \leq N$.
 - (ii) j_1, \dots, j_N are distinct.
 - (iii) $e_{I_n} > 0$ for each $1 \leq n \leq N$.

Here the sets I_1, \dots, I_N correspond to the sequence of poles which we have encountered from performing N contour shifts, and the index j_n corresponds to the variable we have chosen to use to shift the n^{th} contour.

Since the j_i are distinct, the linear forms $L_{0,I_1}, \dots, L_{0,I_N}$ are linearly independent over \mathbb{Q} . We let V_N be their \mathbb{Q} -span and \mathcal{I}_N to be those forms that vanish identically subject to the conditions $L_{0,I_1} = \dots = L_{0,I_N} = 0$. More generally, for $0 \leq n \leq N$ let

$$(8.10) \quad V_n = \text{Span}_{\mathbb{Q}}(L_{0,I_1}(\mathbf{x}), \dots, L_{0,I_n}(\mathbf{x})) \quad \text{and} \quad \mathcal{I}_n = \{I \in \mathcal{S}(2k) : L_{0,I}(\mathbf{x}) \in V_n\},$$

with the conventions that $V_0 = \{0\}$ and $\mathcal{I}_0 = \{\emptyset\}$. Since $j_r \in I_r$ for all r , and j_1, \dots, j_n are distinct integers, if we impose the conditions $\sum_{i \in I_n} x_i = 0$ on the variables x_i , then we may write x_{j_1}, \dots, x_{j_n} as \mathbb{Q} -linear combinations of the other variables. Hence the linear form $L_{0,I}$ becomes a linear form $L_{n,I}$ in the variables x_j , $j \in [2k] \setminus \{j_1, \dots, j_n\}$. Clearly, $L_{0,I} \in V_n$ if and only if $L_{0,I} = 0$ after we have “quotiented” the space of linear forms in the variables x_j with the relations $L_{0,I_1} = \dots = L_{0,I_n} = 0$, if and only if $L_{n,I} = 0$.

Remark. We will show later on that the variables s_1, \dots, s_{2k} can be permuted in a way that allows us to assume that $j_n = 2k - n + 1$ for all n .

Definition 8.1. Let N be a level. The triplet $(\mathbf{I}, \mathbf{h}, d)$ is called a *type of level N* if:

- (a) $\mathbf{I} = (I_1, \dots, I_N)$ is an N -tuple of sets such that $2k - n + 1 \in I_n$ and $e_{I_n} > 0$ for all $n = 1, 2, \dots, N$.
- (b) $\mathbf{h} = (h_{n,I})_{0 \leq n \leq N, I \in \mathcal{S}^*(2k)}$ is a tuple of non-negative integers such that:
 - (i) $h_{n,I} = 0$ for $0 \leq n \leq N$ if $e_I = 0$ (i.e. if $A = 0$ and $\#I = 1$);
 - (ii) $0 = h_{0,I} \leq h_{1,I} \leq \dots \leq h_{N,I}$ for $I \in \mathcal{S}^*(2k)$;
 - (iii) If $I \in \mathcal{I}_n \setminus \mathcal{I}_{n-1}$ for some $n \in [N]$, then $h_{m,I} = h_{n,I}$ for all $m \geq n$.

The integers $h_{n,I}$ will describe the different terms coming up in poles of high order, corresponding to taking many derivatives of different parts of the integrand. (The $h_{N,I}^{\text{th}}$ derivative of $\zeta^{e_I}(1 + L_{N,I}(\mathbf{s}))$ will occur in the integrand of the term we are considering.)
- (c) d is a non-negative integer.

We will further say that the triplet $(\mathbf{I}, \mathbf{h}, d)$ is an *admissible type of level N* if

$$H_N \geq N + d,$$

where

$$H_N = H_N(\mathbf{h}, \mathcal{I}_N, A) := \sum_{I \in \mathcal{I}_N \setminus \{\emptyset\}} (-1)^{\#I} - \sum_{I \in \mathcal{S}(2k) \setminus \mathcal{I}_N} h_{N,I} + (A+1) \sum_{j \in [2k], \{j\} \in \mathcal{I}_N} 1.$$

Remark 8.1. We must have that $H_N \geq N$ if $(\mathbf{I}, \mathbf{h}, d)$ is an admissible type of level N . We will see that the quantity H_N is related to the total order of the poles we have picked up from the first N contour shiftings. We further note that, in the notation of Section 7, it can be written as

$$H_N = \mathcal{A}(V_N) - \sum_{I \in \mathcal{S}(2k) \setminus \mathcal{I}_N} h_{N,I} + (A+1) \sum_{j \in [2k], \{j\} \in \mathcal{I}_N} 1.$$

The data in a type of level N will keep track of all the relevant information on terms we encounter from poles having shifted N contours. Given a type, we can now define the key objects we wish to consider:

Definition 8.2. Let $(\mathbf{I}, \mathbf{h}, d)$ be a type of level N . A function $J : \mathbb{R}_{\geq 2} \rightarrow \mathbb{C}$ is called a *fundamental component of level N and of type $(\mathbf{I}, \mathbf{h}, d)$* if:

- the type $(\mathbf{I}, \mathbf{h}, d)$ is admissible, that is to say, we have $H_N \geq N + d$;
- when $N = 2k$, we have $J(R) = (\log R)^{H_N - N - d - 2kA}$;
- when $N < 2k$, we have

$$J(R) = \frac{(\log R)^{H_N - N - d - 2kA}}{(2i\pi)^{2k-N}} \int \cdots \int_{\substack{\operatorname{Re}(s_j) = \lambda_j / \log R \\ |\operatorname{Im}(s_j)| \leq T \\ 1 \leq j \leq 2k-N}} G(\mathbf{s}) R^{E_N(\mathbf{s})} \\ \times \prod_{I \in \mathcal{S}(2k) \setminus \mathcal{I}_N} (\zeta^{e_I})^{(h_{N,I})} (1 + L_{N,I}(\mathbf{s})) ds_{2k-N} \cdots ds_1$$

where $\lambda_j / \lambda_{j-1} \geq \lambda$,

$$E_N(s_1, \dots, s_{2k-N}) := L_{N,[2k]}(s_1, \dots, s_{2k-N}),$$

and G is a function in the variables s_1, \dots, s_{2k-N} that belongs to the class \mathcal{C}_{2k-N} . Moreover, if we have additionally that $d = 0$, then G is given by

$$G(\mathbf{s}) = F(L_{N,\{1\}}(\mathbf{s}), \dots, L_{N,\{2k\}}(\mathbf{s})).$$

We note that when $d = 0$, we have that G is non-vanishing in Ω_{2k-N} by (8.8) and the preceding discussion.

Definition 8.3. A fundamental component of level N and type $(\mathbf{I}, \mathbf{h}, d)$ is called *irreducible* if either $N = 2k$ or $E_N = 0$. Otherwise, it is called *reducible*.

With the above notation, the integral $I_{g,2k}(R)$ is a reducible fundamental component of level 0 and of type $(\emptyset, \emptyset, 0)$.

If we say that $J(R)$ is a fundamental component of level N , we mean that there exists an admissible type $(\mathbf{I}, \mathbf{h}, d)$ of level N such that $J(R)$ is a fundamental component of level N and type $(\mathbf{I}, \mathbf{h}, d)$.

We begin with a lemma that justifies the terms *irreducible* vs. *reducible*, showing how reducible components are a linear combination of irreducible ones (up to a very small error term). First, we need to introduce a last piece of notation. Notice that if $E_N \neq 0$, then we may uniquely write

$$E_N(\mathbf{x}) = \gamma_1 x_1 + \gamma_2 x_2 + \cdots + \gamma_{j_{N+1}} x_{j_{N+1}}$$

for some $\gamma_j \in \mathbb{Q}$ with $\gamma_{j_{N+1}} \neq 0$. If λ is big enough, then the sign of $\operatorname{Re}(E_N(s))$ throughout the region of integration is constant and equal to the sign of $\gamma_{j_{N+1}}$. The behaviour of reducible fundamental components differs according to this sign:

Lemma 8.4. *Assume the above setup. Let $J(R)$ be a reducible fundamental component of level $N < 2k$ and type $(\mathbf{I}, \mathbf{h}, d)$, and let $\gamma_1, \dots, \gamma_{j_{N+1}}$ be as above. Assume that λ is large enough in terms of $(\mathbf{I}, \mathbf{h}, d)$.*

- (a) *If $\gamma_{j_{N+1}} > 0$, then $J(R)$ is a linear combination of $O(1)$ fundamental components of level $N + 1$ with coefficients of size $O(1)$, up to an error term of size $\ll T^{-1+o(1)}$. Each of these fundamental components has (admissible) type $(\mathbf{I}', \mathbf{h}', d')$ that depends only on $N, \mathbf{I}, \mathbf{h}, d$.*
- (b) *If $\gamma_{j_{N+1}} < 0$, then $J(R) \ll T^{-1+o(1)}$.*

The implied constants depend at most on $(\mathbf{I}, \mathbf{h}, d)$, A and the function G in the definition of J , and are independent of R .

We iterate the above lemma until all the fundamental components we are dealing with are irreducible. For such components, we have the following asymptotic formula.

Lemma 8.5. *Assume the above setup. If $J(R)$ is an irreducible fundamental component, then there is some $c \in \mathbb{C}$ such that*

$$J(R) = c(\log R)^{\mathcal{E}_{k,A}} + O((\log R)^{\mathcal{E}_{k,A}-1}),$$

where we recall that $\mathcal{E}_{k,A} = \max(\binom{2k}{k} - 2k(A+1), -1)$. The implied constant and the constant c are independent of R .

Since $I_{g,2k}(R)$ is a reducible fundamental component of level 0, we apply Lemma 8.4 repeatedly to write it as a linear combination of $O(1)$ irreducible fundamental components, and then estimate these components by Lemma 8.5. This establishes (8.4). We now prove the above two key lemmas.

8.3. Proof of the auxiliary Lemmas 8.4 and 8.5.

Proof of Lemma 8.4. Note that if $E_N \neq 0$, then we must have that either $N = 0$, either $k \geq 2$ or $A \geq 1$: when $N = k = 1$ and $A = 0$, the only $I \subset \{1, 2\}$ with $e_I > 0$ is $I = \{1, 2\}$. But if $x_{\{1,2\}} = 0$, we must have that $E_1 = 0$, a contradiction.

- (a) Here $\gamma_{j_{N+1}} > 0$. For notational simplicity, we make the change of variables

$$s'_j = s_j \ (1 \leq j < j_{N+1}), \ s'_j = s_{j+1} \ (j_{N+1} \leq j < 2k - N), \ s'_{2k-N} = s_{j_{N+1}},$$

which corresponds to a cyclic permutation of the variables $s_{j_{N+1}}, \dots, s_{2k-N}$. We similarly define the linear forms x'_j , using the corresponding permutation of the forms x_j , as well as the parameters λ'_j . We shift the s'_{2k-N} contour to the line $\operatorname{Re}(s'_{2k-N}) = -1/(\log T)^{3/2}$. The contribution of the horizontal integrals is $\ll (\log R)^{O(1)}/T$. Moreover, when $\operatorname{Re}(s'_{2k-N}) = -1/(\log T)^{3/2}$ and $\operatorname{Re}(s'_j) = O(1/\log R)$ for $j < 2k - N$, we have that

$$\operatorname{Re}(E_N(s')) = -\frac{\gamma_{j_{N+1}}}{(\log T)^{3/2}} + O\left(\frac{1}{\log R}\right).$$

It thus follows that the contribution of the integral with $\operatorname{Re}(s'_{2k-N}) = -1/(\log T)^{3/2}$ is $\ll e^{-\sqrt{\log R}}$, say, which is of negligible size. So we need only worry about the poles that the contour shifting introduces.

The poles occur when $L_{N,I_{N+1}}(s') = 0$ for some $I_{N+1} \in \mathcal{S}(2k) \setminus \mathcal{I}_N$ with $e_{I_{N+1}} > 0$ such that the coefficient of s'_{2k-N} in $L_{N,I_{N+1}}$ is non-zero. As we discussed in Section 8.2, imposing the relation $L_{N,I_{N+1}}(x') = 0$ allows us to write x'_{2k-N} as a linear combination of the forms x'_1, \dots, x'_{2k-N-1} , say $x'_{2k-N} = C(x'_1, \dots, x'_{2k-N-1})$. We then define the sets V_{N+1} and \mathcal{I}_{N+1} as in (8.10), and similarly let $E_{N+1} = L_{N+1,[2k]}$.

We need to understand the order of the pole at $s'_{2k-N} = C(s'_1, \dots, s'_{2k-N-1})$. We only look at *generic points* $(s'_1, \dots, s'_{2k-N-1})$: it could be the case that for some measure-zero subset of points, we get a different pole order. For example, for fixed $s_1 \in \mathbb{C}$, the function $s_2 \mapsto s_1/s_2$ has generically a pole of order 1 at $s_2 = 0$, unless $s_1 = 0$, when there is no pole. This reduced pole order however would not affect an integral over s_1 , because it only occurs for a measure-zero set of s_1 values.

With the above discussion in mind, we note that the generic order of the zero of the analytic function

$$\prod_{I \in \mathcal{S}(2k) \setminus \mathcal{I}_{N+1}} (\zeta^{e_I})^{(h_{N,I})} (1 + L_{N,I}(s'))$$

at $s'_{2k-N} = C(s'_1, \dots, s'_{2k-N-1})$ is 0. Indeed, for this product to vanish we must have that $L_{N+1,I}(s') = 0$, which happens non-generically when $I \in \mathcal{S}(2k) \setminus \mathcal{I}_{N+1}$.

Next, let ν be the generic order of the zero of the analytic function

$$(8.11) \quad G(s') \prod_{\substack{I \in \mathcal{I}_{N+1} \setminus \mathcal{I}_N \\ e_I = -1, h_{N,I} \geq 2}} \left(\frac{1}{\zeta} \right)^{(h_{N,I})} (1 + L_{N,I}(s'))$$

at $s'_{2k-N} = C(s'_1, \dots, s'_{2k-N-1})$. If $d = 0$ and $h_{N,I} = 0$ for all $I \in \mathcal{S}^-(2k) \cap (\mathcal{I}_{N+1} \setminus \mathcal{I}_N)$, then the function in (8.11) equals $F(L_{N,\{1\}}(s'), \dots, L_{N,\{2k\}}(s'))$, which does not vanish in Ω_{2k} , so that $\nu = 0$.

From the above discussion, we conclude that the generic order of the pole of the integrand of $J(R)$ at $s'_{2k-N} = C(s'_1, \dots, s'_{2k-N-1})$ is

$$(8.12) \quad \begin{aligned} m &= \sum_{\substack{I \in \mathcal{I}_{N+1} \setminus \mathcal{I}_N \\ \#I = \text{even}}} (h_{N,I} + 1) - \sum_{\substack{I \in \mathcal{I}_{N+1} \setminus \mathcal{I}_N \\ e_I = -1, h_{N,I} = 0}} 1 + \sum_{\substack{1 \leq j \leq 2k \\ \{j\} \in \mathcal{I}_{N+1} \setminus \mathcal{I}_N}} (h_{N,\{j\}} + A) - \nu \\ &= \sum_{I \in \mathcal{I}_{N+1} \setminus \mathcal{I}_N} (h_{N,I} + (-1)^{\#I}) + (A + 1) \sum_{j \in [2k], \{j\} \in \mathcal{I}_{N+1} \setminus \mathcal{I}_N} 1 \\ &\quad - \nu - \sum_{\substack{I \in \mathcal{S}^-(2k) \cap (\mathcal{I}_{N+1} \setminus \mathcal{I}_N) \\ h_{N,I} \geq 2, \#I \geq 3}} (h_{N,I} - 1). \end{aligned}$$

Note that it could be the case that $m \leq 0$, in which case there is no pole contribution to $J(R)$ from the pole with $L_{N,I_{N+1}}(s') = 0$.

Assume, now, that $m \geq 1$. Then, $m + H_N \geq 1 + N + d$ by our assumption that $H_N \geq N + d$. Moreover,

$$(8.13) \quad m + H_N = \sum_{I \in \mathcal{I}_{N+1} \setminus \{\emptyset\}} (-1)^{\#I} - \sum_{I \in \mathcal{S}(2k) \setminus \mathcal{I}_{N+1}} h_{N,I} + (A+1) \sum_{j \in [2k], \{j\} \in \mathcal{I}_{N+1}} 1 - \nu - \sum_{\substack{I \in \mathcal{S}^-(2k) \cap (\mathcal{I}_{N+1} \setminus \mathcal{I}_N) \\ h_{N,I} \geq 1, \#I \geq 3}} (h_{N,I} - 1).$$

In order to continue, we separate two subcases depending on whether $N = 2k - 1$ or $N \leq 2k - 2$.

Case 1 of the proof of Lemma 8.4: $N = 2k - 1$. In this case, we have that $s'_j = L_{2k-1, \{j\}}(s'_1) = a_j s'_1$ for all j , where $a_j \in \mathbb{Q}$. Thus, the only potential pole is at $s'_1 = 0$. If $m \geq 1$, so that there a genuine pole at $s'_1 = 0$, then we obtain an evaluation of $J(R)$ as a finite linear combination of powers of $\log R$ (up to an error term of size $O((\log R)^{O(1)}/T)$), the highest of which has exponent

$$H_{2k-1} + m - 2k - 2kA - d = \sum_{I \in \mathcal{S}^*(2k)} (-1)^{\#I} - \nu - \sum_{\substack{I \in \mathcal{S}^-(2k) \setminus \mathcal{I}_{2k-1} \\ h_{2k-1,I} \geq 2, \#I \geq 3}} (h_{2k-1,I} - 1) - d \leq -1,$$

since $\mathcal{I}_{2k} = \mathcal{S}(2k)$ in this case. We have thus written $J(R)$ as a linear combination of irreducible fundamental components of level $2k$ and suitable type (taking $h_{2k,I} = h_{2k-1,I}$ and $\mathcal{I}_{2k} = \{1\}$), up to a small error term. This proves Lemma 8.4 in this case.

As an amusing remark, we note that the above exponent equals $\mathcal{E}_{k,A}$ only when $A > \frac{1}{2k} \binom{2k}{k} - 1$, $d = 0$, $\nu = 0$, $h_{2k-1,I} \in \{0, 1\}$ for $I \in \mathcal{S}^-(2k) \setminus \mathcal{I}_{2k-1}$, and $G(s) = F(a_1 s, \dots, a_{2k} s)$, in which case the residue is

$$G(0) = A!^{2k} F(0, \dots, 0) = A!^{2k}.$$

Otherwise, these poles contribute towards the error term of $I_{g,2k}(R)$.

Case 2 of the proof of Lemma 8.4: $N \leq 2k - 2$. Then the contribution of the pole $s'_{2k-N} = C(s'_1, \dots, s'_{2k-N-1})$ to $J(R)$ equals

$$\frac{(\log R)^{H_N - N - d - 2kA}}{(2i\pi)^{2k-N-1} m!} \int \cdots \int_{\substack{\text{Re}(s'_j) = \lambda'_j / \log R, \\ |\text{Im}(s'_j)| \leq T \\ 1 \leq j \leq 2k-N-1}} \frac{d^{m-1}}{d(s'_{2k-N})^{m-1}} \Big|_{s'_{2k-N} = C(s'_1, \dots, s'_{2k-N-1})} (Z(\mathbf{s}')) ds'_{2k-N-1} \cdots ds'_1,$$

where

$$Z(\mathbf{s}') := G(\mathbf{s}') R^{E_N(\mathbf{s}')} (s'_{2k-N} - C(s'_1, \dots, s'_{2k-N-1}))^m \prod_{I \in \mathcal{S}(2k) \setminus \mathcal{I}_N} (\zeta^{e_I})^{(h_{N,I})} (1 + L_{N,I}(\mathbf{s}')).$$

Applying the generalized product rule and writing s_j in place of s'_j , we claim that the above integral can be expressed as a finite sum of terms of the form

$$c \cdot \frac{(\log R)^{H_N + m - h - N - 1 - d - 2kA}}{(2i\pi)^{2k-N-1}} \int \cdots \int_{\substack{\text{Re}(s'_j) = \lambda'_j / \log R, |\text{Im}(s'_j)| \leq T \\ 1 \leq j \leq 2k-N-1}} \tilde{G}(\mathbf{s}) R^{E_{N+1}(\mathbf{s})} \times \prod_{I \in \mathcal{S}(2k) \setminus \mathcal{I}_{N+1}} (\zeta^{e_I})^{(h_{N+1,I})} (1 + L_{N+1,I}(\mathbf{s})) ds_{2k-N-1} \cdots ds_1,$$

where:

- $c \ll 1$;
- $h \in \{0, \dots, m-1\}$;
- $h_{N+1,I} \geq h_{N,I}$ with equality if $I \in \mathcal{I}_{N+1} \setminus \{\emptyset\}$;
- $\sum_{I \in \mathcal{S}(2k) \setminus \mathcal{I}_{N+1}} (h_{N+1,I} - h_{N,I}) \leq h$;
- \tilde{G} is in the class \mathcal{C}_{2k-N-1} .

The first four claims are easy to verify, but our claim about \tilde{G} requires some justification. To simplify the notation, we make the change of variables $s'_{2k-N} = \tau + C(s'_1, \dots, s'_{2k-N-1})$. Let $\delta_{N,I}$ denote the coefficient of x'_{2k-N} in the linear form $L_{N,I}$. If $I \in \mathcal{I}_{N+1}$, so that $L_{N+1,I} = 0$, we find that $L_{N,I}(s'_1, \dots, s'_{2k-N}) = \delta_{N,I}\tau$. So, if $I \in \mathcal{I}_{N+1} \setminus \mathcal{I}_N$, then $\delta_{N,I} \neq 0$. Finally, we let $G_1(s'_1, \dots, s'_{2k-N-1}, \tau)$ to be the function $G(s'_1, \dots, s'_{2k-N})$ after our change of variables. If ν_1 denotes the generic order of the zero of $G_1(s'_1, \dots, s'_{2k-N-1}, \tau)$ at $\tau = 0$, then the function \tilde{G} will simply be a linear combination the functions

$$\frac{\partial^j}{\partial \tau^j} \Big|_{\tau=0} (\tau^{-\nu_1} G_1(s'_1, \dots, s'_{2k-N-1}, \tau)) = \frac{j!}{(j + \nu_1)!} \cdot \frac{\partial^{j+\nu_1} G_1}{\partial \tau^{j+\nu_1}}(s'_1, \dots, s'_{2k-N-1}, 0).$$

Since G is in the class \mathcal{C}_{2k-N} , the above functions are in the class \mathcal{C}_{2k-N-1} , which proves our claim about \tilde{G} .

It is straightforward to verify that \mathbf{h}' and \mathbf{I}' satisfy the required properties. Thus it remains to show that there is a suitable d' .

Now, relation (8.13) implies that the exponent of $\log R$ is $H_{N+1} - (N+1) - d' - 2kA$, with

$$d' = d + \nu + \sum_{\substack{I \in \mathcal{S}^-(2k) \cap (\mathcal{I}_{N+1} \setminus \mathcal{I}_N) \\ h_{N,I} \geq 2, \#I \geq 3}} (h_{N,I} - 1) + h - \sum_{I \in \mathcal{S}(2k) \setminus \mathcal{I}_{N+1}} (h_{N+1,I} - h_{N,I}) \geq 0.$$

Moreover, we have that

$$H_{N+1} - d' = H_N - d + m - h \geq N + 1 \implies H_{N+1} \geq d' + N + 1,$$

as needed. Finally, if $d' = 0$, it is easy to check that $G(\mathbf{s}) = F(L_{N+1, \{1\}}(\mathbf{s}), \dots, L_{N+1, \{2k\}}(\mathbf{s}))$.

(b) Here $\gamma_{j_{N+1}} < 0$. We then shift the contours of $s_{2k-N}, s_{2k-N-1}, \dots, s_{j_{N+1}}$ in this order to the lines $\text{Re}(s_j) = \lambda_j / (\log T)^{3/2}$, $j_{N+1} \leq j \leq 2k - N$. If λ is large enough, then we do not encounter any poles and the horizontal lines contribute $\ll (\log R)^{O(1)} / T$ when we make this shift. Finally, when $\text{Re}(s_j) = \lambda_j / (\log T)^{3/2}$ for $j_{N+1} \leq j \leq 2k - N$, and $\text{Re}(s_j) = \lambda^j / \log R$ for $1 \leq j < j_{N+1}$, then we have that

$$\text{Re}(E(\mathbf{s})) = -\frac{|\gamma_{j_{N+1}}| \lambda_{j_{N+1}} (1 + O(1/\lambda))}{(\log T)^{3/2}},$$

so that our integrand is $\ll e^{-\sqrt{\log R}}$ if λ is large enough. We thus find that in this case

$$J(R) \ll T^{-1+o(1)},$$

as needed. □

Remark 8.2. Case 1 is feasible for some choice of I_1, \dots, I_{2k-1} . Indeed, if $k = 1$ and $A \geq 1$, so that $N = 1$, then we note that at least one of the zeta factors must have survived in the numerator after shifting the s_2 contour, and there are none in the denominator, so there is a pole at $s'_1 = 0$. On

the other hand, if $k \geq 2$, then if $I_1 = \{2k-1, 2k\}$, $I_2 = \{2k-1, 2k-2\}$ and $I_3 = \{2k-2, 2k\}$, then $a_{2k} = a_{2k-1} = a_{2k-2} = 0$, and $a_1 \neq 0$. Taking $h_{2k-1,I} = 0$ for all I , we see that

$$m = \sum_{I \in \mathcal{S}(2k) \setminus \mathcal{I}_{2k-1}} (-1)^{\#I} + (A+1) \sum_{j \in [2k], a_j \neq 0} 1 = (A+1) \sum_{j \in [2k], a_j \neq 0} 1 \geq 1,$$

where we used Proposition 7.1(a). Thus we see indeed that there is a genuine pole at $s'_1 = 0$.

It remains to prove the second intermediate step in the proof of (8.4):

Proof of Lemma 8.5. We separate into three cases depending on whether $N = 2k$, $N = 2k-1$ or $N \leq 2k-2$.

Case 1 of the proof of Lemma 8.5: $N = 2k$. Here there is no integral and we have that $J(R) = c(\log R)^{H_{2k}-2k(A+1)-d}$. Since $\mathcal{I}_{2k} = \{I \subset [2k]\}$, we find that $H_{2k} = -1 + 2k(A+1)$, whence $J(R) = (\log R)^{-d-1}$. If $d = 0$ and $\mathcal{E}_{k,A} = -1$, the lemma follows with $c = 1$; otherwise, we take $c = 0$.

Case 2 of the proof of Lemma 8.5: $N = 2k-1$. In this case $J(R)$ is given by a one-dimensional integral over s_1 . Moreover, there exist coefficients $a_j \in \mathbb{Q}$ such that $x_j = L_{2k-1, \{j\}}(x_1) = a_j x_1$ for all j . Therefore the only possible pole of the integrand in $J(R)$ is when $s_1 = 0$. If there is such a pole, then it means that $\{1\} \notin \mathcal{I}_{2k-1}$. In this case, the order of this potential pole, say m , would be given by (8.12) with $N = 2k-1$, $\mathcal{I}_{2k} = \mathcal{S}(2k)$ and ν defined analogously, so that (8.13) implies that

$$(8.14) \quad m + H_{2k-1} - (2k-1) - 2kA = 1 + \sum_{I \in \mathcal{S}^*(2k)} (-1)^{\#I} - \sum_{\substack{I \in \mathcal{S}^-(2k) \setminus \mathcal{I}_{2k-1} \\ \#I \geq 3, h_{2k-1,I} \geq 2}} (h_{2k-1,I} - 1) - \nu \leq 0,$$

First, let us assume $m \geq 1$ (i.e. there is a genuine pole at $s_1 = 0$). We find that $H_{2k-1} + 1 - 2k(A+1) \leq -1$. We then move the line of integration of $s_1 = \sigma_1 + it_1$ to the contour $\sigma_1 = 1/(\log(2+|t_1|))^{3/2}$, $|t_1| \leq T$. No poles are encountered and the horizontal integrals contribute $\ll (\log R)^{O(1)}/T$. Moreover, the integral converges fast enough now (even when $A = 0$) that we may remove the condition $|t_1| \leq T$ at the cost of an error term of size $\ll (\log R)^{O(1)}/T$. We thus conclude that

$$J(R) = c \cdot (\log R)^{H_{2k-1}-2k(A+1)+1-d} + O(T^{-1+o(1)}),$$

where

$$c = \int_{\sigma_1=1/(\log(2+|t_1|))^{3/2}} G(s_1) \prod_{I \in \mathcal{S}(2k) \setminus \mathcal{I}_{2k-1}} (\zeta^{e_I})^{(h_{2k-1,I})} (1 + a_I s_1) ds_1$$

is some constant. This contributes towards the error term if $m \geq 2$, $d \geq 1$ or $A \leq \frac{1}{2k} \binom{2k}{k} - 1$, and towards the main term if $m = 1$, $d = 0$, $A > \frac{1}{2k} \binom{2k}{k} - 1$ and $h_{2k-1,I} \in \{0, 1\}$ for each $I \in \mathcal{S}^-(2k)$, in which case $H_{2k-1} - 2k(A+1) + 1 = -1$ by (8.14).

Alternatively, assume that $m \leq 0$, so that there is no pole at $s_1 = 0$. We move s_1 to the line $\text{Re}(s_1) = 0$. The horizontal lines contribute $\ll (\log R)^{O(1)}/T$. Furthermore, we note that we may extend the range of integration to all $s_1 \in \mathbb{C}$ with $\text{Re}(s_1) = 0$ at the cost of an error term of size $\ll (\log R)^{O(1)}/T$. Consequently,

$$J(R) = c \cdot (\log R)^{H_{2k-1}+1-2k(A+1)-d} + O((\log R)^{O(1)}/T),$$

where

$$c = \frac{1}{2\pi} \int_{-\infty}^{\infty} G(it) \prod_{I \in \mathcal{S}(2k) \setminus \mathcal{I}_{2k-1}} (\zeta^{e_I})^{(h_{2k-1,I})} (1 + ia_I t) dt$$

with $a_I := \sum_{j \in I} a_j$. The power of $\log R$ is

$$\begin{aligned} H_{2k-1} + 1 - 2k(A+1) - d &= 1 + \sum_{I \in \mathcal{S}_{2k-1} \setminus \{0\}} (-1)^{\#I} - \sum_{I \in \mathcal{S}(2k) \setminus \mathcal{I}_{2k-1}} h_{2k-1,I} \\ &\quad - (A+1) \cdot \#\{1 \leq j \leq 2k : a_j \neq 0\} - d \\ &= 1 + \mathcal{A}(V_{2k-1}) - \sum_{I \in \mathcal{S}(2k) \setminus \mathcal{I}_{2k-1}} h_{2k-1,I} \\ &\quad - (A+1) \cdot \#\{1 \leq j \leq 2k : a_j \neq 0\} - d \end{aligned}$$

in the notation of Proposition 7.1. Clearly, this is maximized when $h_{2k-1,I} = 0$ for all $I \in \mathcal{S}(2k) \setminus \mathcal{I}_{2k-1}$ and $d = 0$, in which case

$$(8.15) \quad G(s) = F(a_1 s, \dots, a_{2k} s) = \frac{P(a_1 s, \dots, a_{2k} s)}{\prod_{j=1}^{2k} (a_j s \zeta(1 + a_j s))^{A+1}} + O\left(\frac{(\log(2 + |s|))^{O(1)}}{(1 + |s|)^{(2k-1)(A+1)}}\right)$$

when $\operatorname{Re}(s) = 0$, by (6.6).

Now, if $a_j = 0$ for some $j \in [2k]$, then $\mathcal{A}(V_{2k-1}) = -1$, by Proposition 7.1(a). Note that there is at least one j such that $a_j \neq 0$; otherwise, the dimension of V_{2k-1} would be $2k$, as it would contain the independent forms s_1, \dots, s_{2k} , which is a contradiction. We conclude that the power of $\log R$ is $\leq -(A+1) \cdot \#\{1 \leq j \leq 2k : a_j \neq 0\} \leq -A-1$. Consequently,

$$J(R) \ll (\log R)^{-A-1}$$

in this case, which contributes towards the error term (i.e. $c = 0$ in this case).

Finally, assume that $a_j \neq 0$ for all $j \in [2k]$. Since $\dim(V_{2k-1}) = 2k-1$, Proposition 7.1(b) implies that the power of $\log R$ is

$$H_{2k-1} + 1 - 2k(A+1) - d \leq \mathcal{A}(V_{2k-1}) - \dim(V_{2k-1}) - 2kA \leq \binom{2k}{k} - 2kA,$$

with the second inequality being an equality when half of the a_j 's equal $+1$ and the other half -1 .

Even though this not needed for the proof, we remark that when $a_j \neq 0$ for all j , we can give an asymptotic formula for $J(R)$. For simplicity, let us assume that $a_j = 1$ for $j \leq k$ and $a_j = -1$ for $j > k$. Then $a_I = \#(I \cap [1, k]) - \#(I \cap (k, 2k])$, which has the same parity as $\#I$. In particular, $\mathcal{I}^-(2k) \cap \mathcal{I}_{2k-1} = \emptyset$, so that $h_{2k-1,I} = 0$ for all $I \in \mathcal{S}^-(2k)$. Moreover, given $\ell \in \mathbb{Z}$, we have that $a_I = \ell$ for exactly $\binom{2k}{k+|\ell|}$ sets $I \subset [2k]$. Therefore

$$\begin{aligned} J(R) &= \frac{(\log R)^{\binom{2k}{k}-2kA}}{2\pi} \int_{-\infty}^{\infty} \frac{P(it, -it, \dots, it, -it)}{t^{2k(A+1)}} \cdot \frac{\prod_{\ell \text{ even}, \geq 2} |\zeta(1 + i\ell t)|^{2\binom{2k}{k+\ell}}}{\prod_{\ell \text{ odd}, \geq 1} |\zeta(1 + i\ell t)|^{2\binom{2k}{k+\ell}}} dt \\ &\quad + O(T^{-1+o(1)}). \end{aligned}$$

This completes the study of Case 2.

Case 3 of the proof of Lemma 8.5: $N \leq 2k-2$. We shift the contours of $s_{2k-N}, s_{2k-N-1}, \dots, s_1$ in this order to the lines $\operatorname{Re}(s_j) = \lambda^j / (\log T)^{3/2}$, $1 \leq j \leq 2k-N$. If λ is large enough in terms of k (but independently of R), then the functions $\operatorname{Re}(L_{N,I}(s))$ with $I \in \mathcal{S}(2k) \setminus \mathcal{I}_N$ have constant sign in the entire domain where the contour shifting is performed, so that no poles are encountered.

The horizontal lines contribute $\ll (\log R)^{O(1)}/T$. Finally, we note that the integrand on the new lines of integration is $\ll (\log \log R)^{O(1)}/[(1 + |s_1|) \cdots (1 + |s_{2k-N}|)]$, by (8.8) and our assumption that G is in the class \mathcal{C}_{2k-N} . We thus find that

$$J(R) \ll (\log R)^{H_N - N - 2kA - d} (\log \log R)^{O(1)}$$

in this case. We need to understand the power of $\log R$. Firstly, note that

$$H_N - N - 2kA - d \leq 2k - N + \sum_{I \in \mathcal{I}_N \setminus \{\emptyset\}} (-1)^{\#I} - (A+1) \cdot \#\{1 \leq j \leq 2k : \{j\} \notin \mathcal{I}_N\}.$$

To continue, we separate two cases.

If there is $\{j\} \in \mathcal{I}_N$, then $\sum_{I \in \mathcal{I}_N \setminus \{\emptyset\}} (-1)^{\#I} = -1$ by Proposition 7.1(a). Since $\dim(V_N) = N$ by construction, there are $\leq N$ integers j with $\{j\} \in \mathcal{I}_N$. The power of $\log R$ is thus

$$\leq 2k - N - 1 - (A+1)(2k - N) = -1 - A(2k - N) \leq -1 - 2 \cdot \mathbf{1}_{A \geq 1},$$

and we can see that $J(R)$ satisfies the conclusion of the lemma with no main term (i.e. $c = 0$). So assume that there is no $\{j\} \in \mathcal{I}_N$. Then we find that

$$H_N - N - 2kA \leq 2k - N + \sum_{I \in \mathcal{I}_N \setminus \{\emptyset\}} (-1)^{\#I} - 2k(A+1) \leq \binom{2k}{k} - 2k(A+1) - 2,$$

by Proposition 7.1(c), which again means that the lemma holds with $c = 0$. \square

8.4. Dyadic intervals. We conclude this section with a brief explanation of the proof of (8.2). We have that

$$\mathcal{M}_{\tilde{f}_0, 2k}(R) = \mathcal{M}_{\tilde{h}, 2k}(R) + O\left(\frac{1}{\log R}\right),$$

where $\tilde{h}(x) = h(x) - h(x + \frac{\log 2}{\log R})$, by the argument leading to (6.4). We then note that

$$(\widehat{\tilde{h}})_R(s) = \widehat{h}_R(s)(1 - 2^{-s}),$$

so that Perron's inversion formula and relation (6.2) imply that

$$\mathcal{M}_{\tilde{f}_0, 2k}(R) = \frac{1}{(2i\pi)^{2k}} \int \cdots \int_{\substack{\operatorname{Re}(s_j) = \lambda^j / \log R \\ 1 \leq j \leq 2k}} D(\mathbf{s}) \left(\prod_{j=1}^{2k} \widehat{h}_R(s_j)(1 - 2^{-s_j}) \right) ds_{2k} \cdots ds_1 + O\left(\frac{1}{\log R}\right),$$

where λ and T are as before. We thus find that

$$(8.16) \quad \begin{aligned} \mathcal{M}_{\tilde{f}_0, 2k}(R) &= \frac{1}{(2i\pi)^{2k}} \int \cdots \int_{\substack{\operatorname{Re}(s_j) = \lambda^j / \log R \\ 1 \leq j \leq 2k}} \tilde{F}(\mathbf{s}) R^{s_1 + \cdots + s_{2k}} \prod_{I \in \mathcal{S}^*(2k)} \zeta(1 + s_I)^{(-1)^{\#I}} ds_1 \cdots ds_{2k} \\ &\quad + O\left(\frac{1}{\log R}\right), \end{aligned}$$

where

$$\tilde{F}(\mathbf{s}) := P(\mathbf{s}) \prod_{j=1}^{2k} \frac{\widehat{h}_R(s_j)(1 - 2^{-s_j})}{R^{s_j}}$$

and P is defined as above. The function \tilde{F} is in the class \mathcal{C}_{2k} , since the factor $1 - 2^{-s_j}$ annihilates the pole of $\widehat{h}_R(s_j)$ at $s_j = 0$. We thus see that the above integral has the same shape as the integral

$I_{g,2k}(R)$ with $A = 0$, with the difference that $e_I = -1$ when $\#I = 1$. We thus follow the argument leading to (8.1) when $A = 0$ with the obvious modifications. The only difference is that in the analogue of (8.12) we have instead

$$\begin{aligned} m &= -\nu + \sum_{\substack{I \in \mathcal{I}_{N+1} \setminus \mathcal{I}_N \\ \#I = \text{even}}} (h_{N,I} + 1) - \sum_{\substack{I \in \mathcal{I}_{N+1} \setminus \mathcal{I}_N \\ \#I = \text{odd}, h_{N,I} = 0}} 1 \\ &= -\nu + \sum_{I \in \mathcal{I}_{N+1} \setminus \mathcal{I}_N} (h_{N,I} + (-1)^{\#I}) - \sum_{\substack{I \in \mathcal{S}^-(2k) \cap (\mathcal{I}_{N+1} \setminus \mathcal{I}_N) \\ h_{N,I} \geq 2, \#I \geq 3}} (h_{N,I} - 1), \end{aligned}$$

with ν defined as in the proof of Lemma 8.4. We thus find that m has the same expression as when $A = -1$, and relation (8.2) follows by the proof of (8.4) when $A = -1$. An important remark is that when $A = -1$ there is a power of $(\log R)^{-2k}$ in the denominator of the integrand of $I_{g,2k}(R)$ that is not present in the denominator of the right hand side of (8.16).

9. LOWER BOUNDS

In this section we complete our proof of Theorem 1.3 by showing that, for fixed $k \in \mathbb{Z}_{\geq 1}$, $A \in \mathbb{Z}_{\geq 0}$ and $\epsilon > 0$, there are positive constants $c'_{k,A} > 0$ and $c''_k > 0$ such that

$$(9.1) \quad \mathcal{M}_{f_A,2k}(R) \geq c'_{k,A} (\log R)^{\mathcal{E}_{k,A} - \epsilon} + O_\epsilon((\log R)^{\mathcal{E}_{k,A} - 1}),$$

and

$$(9.2) \quad \mathcal{M}_{\tilde{f}_0,2k}(R) \geq c''_k (\log R)^{D_{k,0} - \epsilon} + O_\epsilon((\log R)^{D_{k,0} - 1}),$$

where $\tilde{f}_0(x) = f_0(x) - f_0(x + \frac{\log 2}{\log R})$, as before. Evidently, this completes the proof of Theorem 1.3, since if the constants in the leading terms in (8.1) or (8.2) were 0, we would obtain a contradiction to the above lower bounds by letting $R \rightarrow \infty$.

As in the previous section, there is some smooth function h such that $h(x) = 1$ for $x \leq 1 - 1/(\log R)^C$ and $h(x) = 0$ for $x \geq 1$, with $C = (2k - 1)2^{2k+1} + 2k + 2$, so that

$$\mathcal{M}_{f_0,2k}(R) = \mathcal{M}_{h,2k}(R) + O\left(\frac{1}{\log R}\right) \quad \text{and} \quad \mathcal{M}_{\tilde{f}_0,2k}(R) = \mathcal{M}_{\tilde{h},2k}(R) + O\left(\frac{1}{\log R}\right),$$

where $\tilde{h}(x) = h(x) - h(x + \frac{\log 2}{\log R})$. So, it suffices to prove that

$$(9.3) \quad \mathcal{M}_{g,2k}(R) \geq c'_{k,A} (\log R)^{\mathcal{E}_{k,A} - \epsilon} + O((\log R)^{\mathcal{E}_{k,A} - 1}),$$

where $g \in \{h, \tilde{h}\}$ when $A = 0$, and $g = f_A$ when $A \geq 1$.

Positivity is a key to this proof: we will consider the sum restricted to those integers with a convenient prime factorization, which clearly provides a lower bound. The fact that these integers have a convenient prime factorization means the corresponding sum is technically easier to analyze.

9.1. First manipulations. To ease notation, let

$$\Pi_R := \prod_{p \leq R} \left(1 - \frac{1}{p}\right)$$

We start by observing that

$$\mathcal{M}_{g,2k}(R) = \Pi_R \sum_{P^+(n) \leq R} \frac{1}{n} \left(\sum_{d|n} \mu(d) g \left(\frac{\log d}{\log R} \right) \right)^{2k},$$

since $\text{supp}(g) \subset (-\infty, 1]$. So (9.1) follows immediately when $A > -1 + \frac{1}{2k} \binom{2k}{k}$ by noticing that $\mathcal{E}_{k,A} = -1$ in this case and that we always have that $\mathcal{M}_{g,2k}(R) \geq \Pi_R \gg 1/\log R$.

For the rest of the proof, we assume that $A \leq -1 + \frac{1}{2k} \binom{2k}{k}$, so that

$$\mathcal{E}_{k,A} = \binom{2k}{k} - 2k(A+1) \geq 0.$$

Let $q_1 < q_2 < \dots$ be the sequence of all prime numbers and set

$$Q = \prod_{j=1}^{A+1} q_j.$$

If $A \geq 1$, or if $A = 0$ and $g = h$, then we restrict our attention to integers of the form Qn with $P^-(n) > Q$, so that

$$\mathcal{M}_{g,2k}(R) \geq \frac{\Pi_R}{Q} \sum_{p|n \Rightarrow q_{A+1} < p \leq R} \frac{1}{n} \left(\sum_{J \subset [A+1]} (-1)^{\#J} \sum_{d|n} \mu(d) g \left(\frac{\log(q_J d)}{\log R} \right) \right)^{2k},$$

where $q_J := \prod_{j \in J} q_j$. We define

$$(9.4) \quad w(x) := \sum_{J \subset [A+1]} (-1)^{\#J} g \left(x + \frac{\log q_J}{\log R} \right),$$

so that

$$(9.5) \quad \mathcal{M}_{g,2k}(R) \geq \frac{\Pi_R}{Q} \sum_{p|n \Rightarrow q_{A+1} < p \leq R} \frac{1}{n} \left(\sum_{d|n} \mu(d) w \left(\frac{\log d}{\log R} \right) \right)^{2k}.$$

We further note that $w = \tilde{h}$ when $A = 0$, so that the right hand side of (9.5) is a trivial lower bound for $\mathcal{M}_{\tilde{h},2k}(R)$. Therefore, relation (9.3) will follow in all cases if we can show that

$$(9.6) \quad \begin{aligned} W &:= \Pi_R \sum_{p|n \Rightarrow q_{A+1} < p \leq R} \frac{1}{n} \left(\sum_{d|n} \mu(d) w \left(\frac{\log d}{\log R} \right) \right)^{2k} \\ &\geq c''_{k,A} (\log R)^{\mathcal{E}_{k,A} - \epsilon} + O_\epsilon((\log R)^{\mathcal{E}_{k,A} - 1}) \end{aligned}$$

for some $c''_{k,A} > 0$, where w is defined by (9.4) with $g = h$ if $A = 0$ and $g = f_A$ if $A \geq 1$.

Next, observe that

$$(9.7) \quad \widehat{w}_R(s) = (\log R) \sum_{J \subset [A+1]} (-1)^{\#J} \int_{-\infty}^{\infty} g \left(u + \frac{\log q_J}{\log R} \right) R^{su} du = \widehat{g}_R(s) \prod_{j=1}^{A+1} (1 - q_j^{-s}).$$

In particular, we see that \widehat{w}_R has an analytic continuation to \mathbb{C} and it satisfies the bound

$$(9.8) \quad \widehat{w}_R(s) \ll \frac{R^{\operatorname{Re}(s)}}{(|s|+1)(\log R)^A} \quad (\operatorname{Re}(s) \geq -1),$$

which follows by the definition of f_A when $A \geq 1$ and by (6.5) otherwise. Finally, we note that we also have the bound

$$(9.9) \quad \widehat{w}_R(s) \ll \frac{R^{\operatorname{Re}(s)}(\log R)^{O(1)}}{1+|s|^2} \quad (\operatorname{Re}(s) \geq 1/\log R),$$

which follows from (6.2) when $A = 0$. (This bound can be shown to hold in a larger range, but the above range is good enough for our purposes.)

Before we apply Perron's inversion formula to write the right hand side of (9.5) in terms of \widehat{w}_R , we use positivity again to focus on integers n of a certain convenient form. We set

$$y = \exp\{(\log R)^{1-\epsilon'}\} \quad \text{and} \quad Y = \exp\{(\log R)^{1-\epsilon'/2}\},$$

where $\epsilon' > 0$ will be taken to be small enough in terms of ϵ , and write

$$\mathcal{N} = \{n \in \mathbb{Z}_{\geq 1} : p|n \Rightarrow q_{A+1} < p \leq y\}.$$

We then focus on integers of the form $n = mp_1 \cdots p_k$, where $m \in \mathcal{N}$ with $m \leq Y$, and p_1, \dots, p_k are distinct primes from the interval $(R^{1/2}, R]$. For such an integer n , if $d|n$, then either $d = d'$ or $d = d'p_\ell$, for some $d'|m$ and some $\ell \in \{1, \dots, k\}$. So, we conclude that

$$\begin{aligned} W &\geq \frac{\Pi_R}{k!} \sum_{\substack{m \in \mathcal{N} \\ m \leq Y}} \sum_{\substack{\sqrt{R} < p_1, \dots, p_k \leq R \\ \text{distinct primes}}} \frac{1}{mp_1 \cdots p_k} \\ &\quad \times \left(\sum_{\ell=1}^k \sum_{d|m} \mu(d) w \left(\frac{\log(p_\ell d)}{\log R} \right) - \sum_{d|m} \mu(d) w \left(\frac{\log d}{\log R} \right) \right)^{2k}. \end{aligned}$$

Note that the condition that $m \leq Y = R^{o(1)}$ certainly implies that $d < R/(2q_{A+1})$ for all divisors d of m (since $q_{A+1} \ll_A 1$), so that

$$\begin{aligned} \sum_{d|m} \mu(d) w \left(\frac{\log d}{\log R} \right) &= \sum_{J \subset [A+1]} (-1)^{\#J} \sum_{d|m} \mu(d) \left(1 - \frac{\log(q_J d)}{\log R} \right)^A \\ &= \sum_{d|Qm} \mu(d) \left(1 - \frac{\log d}{\log R} \right)^A = 0 \end{aligned}$$

by observing that $\Delta^{(A+1)}(1-x)^A = 0$ by (1.7), and by applying (1.6) with $r = A+1$, since $\omega(Qm) \geq \omega(Q) \geq A+1$ here. We thus conclude that

$$\begin{aligned} W &\geq \frac{\Pi_R}{k!} \sum_{\substack{m \in \mathcal{N} \\ m \leq Y}} \sum_{\substack{\sqrt{R} < p_1, \dots, p_k \leq R \\ \text{distinct primes}}} \frac{1}{mp_1 \cdots p_k} \left(\sum_{\ell=1}^k \sum_{d|m} \mu(d) w \left(\frac{\log(p_\ell d)}{\log R} \right) \right)^{2k} \\ &\geq \frac{\Pi_R}{k!(\log R)^k} \sum_{\substack{m \in \mathcal{N} \\ m \leq Y}} \sum_{\substack{\sqrt{R} < p_1, \dots, p_k \leq R \\ \text{distinct primes}}} \frac{\prod_{j=1}^k \log p_j}{mp_1 \cdots p_k} \left(\sum_{\ell=1}^k \sum_{d|m} \mu(d) w \left(\frac{\log(p_\ell d)}{\log R} \right) \right)^{2k}. \end{aligned}$$

We note that

$$\sum_{n \in \mathcal{N}, n > Y} \frac{\tau_r(n)}{n} \leq Y^{-1/\log y} \sum_{n \in \mathcal{N}} \frac{\tau_r(n)}{n^{1-1/\log y}} \ll \frac{1}{(\log R)^B},$$

for any fixed $B \geq 1$ and $r \in \mathbb{Z}_{\geq 1}$. Therefore, we may drop the conditions that $m \leq Y$ and that the p_ℓ 's are distinct at the cost of an admissible error term, finding that

$$W \geq \frac{\Pi_R}{k!(\log R)^k} \sum_{m \in \mathcal{N}} \sum_{\sqrt{R} < p_1, \dots, p_k \leq R} \frac{\prod_{j=1}^k \log p_j}{m p_1 \cdots p_k} \left(\sum_{\ell=1}^k \sum_{d|m} \mu(d) w \left(\frac{\log(p_\ell d)}{\log R} \right) \right)^{2k} + O \left(\frac{1}{(\log R)^{100}} \right).$$

Next, we expand the $2k$ -th power as follows:

$$\left(\sum_{\ell=1}^k \sum_{d|m} \mu(d) w \left(\frac{\log(p_\ell d)}{\log R} \right) \right)^{2k} = \sum_{\substack{J_1 \cup \dots \cup J_k = [2k] \\ J_i \cap J_j = \emptyset \text{ for } i \neq j}} \prod_{\ell=1}^k \prod_{j \in J_\ell} \sum_{d_j | m} \mu(d_j) w \left(\frac{\log(p_\ell d_j)}{\log R} \right),$$

with the convention that if $J_\ell = \emptyset$, then the corresponding factor equals 1. Clearly, the summands corresponding to those k -tuples $\mathbf{J} = (J_1, \dots, J_k)$ all of whose components have an even number of terms are non-negative. We write \mathcal{J} for the set of k -tuples \mathbf{J} such that (J_1, \dots, J_k) is a partition of $[2k]$ and either $\#J_\ell = 2$ for all ℓ , or there is some ℓ such that $\#J_\ell$ is an odd number. Then

$$\left(\sum_{\ell=1}^k \sum_{d|m} \mu(d) w \left(\frac{\log(p_\ell d)}{\log R} \right) \right)^{2k} \geq \sum_{\mathbf{J} \in \mathcal{J}} \prod_{\ell=1}^k \prod_{j \in J_\ell} \sum_{d_j | m} \mu(d_j) w \left(\frac{\log(p_\ell d_j)}{\log R} \right).$$

Moreover, if $D \in \mathcal{N}$, then

$$\sum_{\substack{m \in \mathcal{N} \\ D|m}} \frac{1}{m} = \frac{1}{D} \prod_{q_{A+1} < p \leq y} \left(1 - \frac{1}{p} \right)^{-1} = \frac{c_1 \prod_{p \leq y} (1 - 1/p)^{-1}}{D}$$

with $c_1 = \prod_{p \leq q_{A+1}} (1 - 1/p) \asymp 1$. So, we conclude that

$$(9.10) \quad W \geq \frac{c_1 \prod_{y < p \leq R} (1 - 1/p)}{k!(\log R)^k} \sum_{\mathbf{J} \in \mathcal{J}} \sum_{d_1, \dots, d_{2k} \in \mathcal{N}} \frac{\mu(d_1) \cdots \mu(d_{2k})}{[d_1, \dots, d_{2k}]} \times \prod_{\ell=1}^k \sum_{\sqrt{R} < p_\ell \leq R} \frac{\log p_\ell}{p_\ell} \prod_{j \in J_\ell} w \left(\frac{\log(p_\ell d_j)}{\log R} \right) - O \left(\frac{1}{(\log R)^{100}} \right).$$

For the convenience of notation, set

$$W(\mathbf{J}) = \frac{1}{(\log R)^k} \sum_{d_1, \dots, d_{2k} \in \mathcal{N}} \frac{\mu(d_1) \cdots \mu(d_{2k})}{[d_1, \dots, d_{2k}]} \prod_{\ell=1}^k \sum_{\sqrt{R} < p_\ell \leq R} \frac{\log p_\ell}{p_\ell} \prod_{j \in J_\ell} w \left(\frac{\log(p_\ell d_j)}{\log R} \right),$$

so that

$$W \geq \frac{c_1 \prod_{y < p \leq R} (1 - 1/p)}{k!} \sum_{\mathbf{J} \in \mathcal{J}} W(\mathbf{J}) - O \left(\frac{1}{(\log R)^{100}} \right).$$

We will show that the dominant contribution comes from the terms \mathbf{J} with $\#J_\ell = 2$ for all ℓ .

9.2. Mellin transformation. Fix a choice of sets $\mathbf{J} = (J_1, \dots, J_k) \in \mathcal{J}$ and let

$$\mathcal{L} = \{1 \leq \ell \leq k : J_\ell \neq \emptyset\}.$$

For $\ell \notin \mathcal{L}$, the sum over p_ℓ is

$$\sum_{\sqrt{R} < p_\ell \leq R} \frac{\log p_\ell}{p_\ell} = \frac{\log R}{2} + O\left(e^{-c\sqrt{\log R}}\right).$$

So

$$W(\mathbf{J}) = \frac{2^{\#\mathcal{L}-k}}{(\log R)^{\#\mathcal{L}}} \sum_{d_1, \dots, d_{2k} \in \mathcal{N}} \frac{\mu(d_1) \cdots \mu(d_{2k})}{[d_1, \dots, d_{2k}]} \prod_{\ell \in \mathcal{L}} \sum_{p_\ell > \sqrt{R}} \frac{\log p_\ell}{p_\ell} \prod_{j \in J_\ell} w\left(\frac{\log(p_\ell d_j)}{\log R}\right) + O\left(\frac{1}{(\log R)^{100}}\right),$$

where the condition that $p_\ell \leq R$ was dropped because it is implied by the fact that $\text{supp}(w) \subset (-\infty, 1]$.

Next, we use Perron's formula $2k$ times to write each appearance of w as an integral of \widehat{w}_R . We thus find that

$$(9.11) \quad W(\mathbf{J}) = \frac{2^{\#\mathcal{L}-k} (\log R)^{-\#\mathcal{L}}}{(2\pi i)^{2k}} \int \cdots \int_{\substack{\text{Re}(s_j)=1/\log R \\ 1 \leq j \leq 2k}} \sum_{d_1, \dots, d_{2k} \in \mathcal{N}} \frac{\prod_{j=1}^{2k} \mu(d_j) d_j^{-s_j}}{[d_1, \dots, d_{2k}]} \left(\prod_{\ell \in \mathcal{L}} \sum_{p_\ell > \sqrt{R}} \frac{\log p_\ell}{p_\ell^{1+s_{J_\ell}}} \right) \times \left(\prod_{j=1}^{2k} \widehat{w}_R(s_j) \right) ds_1 \cdots ds_{2k} + O\left(\frac{1}{\log R}\right),$$

with the notational convention that $s_J := \sum_{j \in J} s_j$. By possibly re-indexing the variables s_1, \dots, s_{2k} , we may assume that $\mathcal{L} = \{1, \dots, L\}$, where $L = \#\mathcal{L}$, and that $\max J_\ell = 2k - L + \ell$ for all $\ell \in \{1, \dots, L\}$. We want to move the variables $s_{2k-L+1}, \dots, s_{2k}$ to the left. First, we need some bounds on the sum over p_ℓ . We note that

$$\sum_{p > \sqrt{R}} \frac{\log p}{p^{1+s}} = -\frac{\zeta'}{\zeta}(1+s) + O(1) - \sum_{p \leq R^{1/2}} \frac{\log p}{p^{1+s}}$$

for $\text{Re}(s) \geq -1/3$. Using standard bounds on the Riemann zeta function (see, for example, Titchmarsh [26, Theorem 3.11]), we find that

$$\begin{aligned} \sum_{p > \sqrt{R}} \frac{\log p}{p^{1+s}} &\ll \frac{1}{|s|} + \log(2 + |t|) + R^{\max\{0, -\sigma\}/2} \sum_{p \leq \sqrt{R}} \frac{\log p}{p} \\ &\ll R^{\max\{0, -\sigma\}/2} \log(R + |t|), \end{aligned}$$

where $s = \sigma + it$, as usual. Moreover, note that if $s_j \geq -1/\log y$ for all $j \in \{1, \dots, 2k\}$, then

$$\begin{aligned} \left| \sum_{d_1, \dots, d_{2k} \in \mathcal{N}} \frac{\prod_{j=1}^{2k} \mu(d_j) d_j^{-s_j}}{[d_1, \dots, d_{2k}]} \right| &\leq \sum_{d_1, \dots, d_{2k} \in \mathcal{N}} \frac{\prod_{j=1}^{2k} \mu^2(d_j) d_j^{1/\log y}}{[d_1, \dots, d_{2k}]} \\ &\leq \prod_{p \leq y} \left(1 + \frac{p^{1/\log y} (4^k - 1)}{p} \right) \ll (\log y)^{4^k - 1}, \end{aligned}$$

using the estimate $p^{1/\log y} = 1 + O(\log p / \log y)$ for $p \leq y$.

We are now ready to move the variables $s_{2k-L+1}, \dots, s_{2k}$ in (9.11) to the left. First, we move the variable s_{2k} to the line $\operatorname{Re}(s_{2k}) = -1/\log y$. (Here we can use (9.9) to justify the convergence required for this manoeuvre.) We pick up a simple pole from the sum over p_{2k} when $s_{J_L} = 0$. The integrand when $\operatorname{Re}(s_{2k}) = -1/\log y$ is

$$\ll (\log y)^{4^k-1} (\log R)^{L-1} \log(R+|t|) R^{1/(2\log y)} \prod_{j=1}^{2k} |\widehat{w}_R(s_j)| \ll \frac{(\log R)^{O(1)} \log(2+|t|) R^{-1/(2\log y)}}{(|s_1|^2+1) \cdots (|s_{2k}|^2+1)}$$

by (9.9). So we find that

$$\begin{aligned} W(\mathbf{J}) &= \frac{2^{L-k} (\log R)^{-L}}{(2\pi i)^{2k-1}} \int \cdots \int_{\substack{\operatorname{Re}(s_j)=1/\log R \\ 1 \leq j \leq 2k-1 \\ s_{J_L}=0}} \sum_{d_1, \dots, d_{2k} \in \mathcal{N}} \frac{\prod_{j=1}^{2k} \mu(d_j) d_j^{-s_j}}{[d_1, \dots, d_{2k}]} \left(\prod_{\ell=1}^{L-1} \sum_{p_\ell > R^{1/2}} \frac{\log p_\ell}{p_\ell^{1+s_{J_\ell}}} \right) \\ &\quad \times \left(\prod_{j=1}^{2k} \widehat{w}_R(s_j) \right) ds_1 \cdots ds_{2k-1} + O\left(\frac{1}{(\log R)^{100}}\right). \end{aligned}$$

Now the product $\prod_{\ell=1}^{L-1} \sum_{p_\ell > R^{1/2}} (\log p_\ell) p_\ell^{-1-s_{J_\ell}}$ doesn't depend on the variables $s_j \in J_L \setminus \{2k\}$, and so we encounter no poles if we move all of these variables to the lines $\operatorname{Re}(s_j) = 0$. Having done this, by (9.8) the growth of $\widehat{w}_R(s_j)$ only depends weakly on R .

Then we repeat the same argument by moving s_{2k-1} to the left, then s_{2k-2} , and so on and so forth, until all the sums over primes have been removed and replaced by contributions coming from poles. Writing $s_j = it_j$, we conclude that

$$\begin{aligned} W(\mathbf{J}) &= \frac{2^{L-k} (\log R)^{-L}}{(2\pi)^{2k-L}} \int \cdots \int_{\substack{t_1, \dots, t_{2k} \in \mathbb{R} \\ t_{J_\ell}=0 \ (1 \leq \ell \leq L)}} \sum_{d_1, \dots, d_{2k} \in \mathcal{N}} \frac{\prod_{j=1}^{2k} \mu(d_j) d_j^{-it_j}}{[d_1, \dots, d_{2k}]} \left(\prod_{j=1}^{2k} \widehat{w}_R(it_j) \right) dt_1 \cdots dt_{2k-L} \\ (9.12) \quad &\quad + O\left(\frac{1}{(\log R)^{100}}\right). \end{aligned}$$

We note that, for any t_1, \dots, t_{2k} , we have

$$\sum_{d_1, \dots, d_{2k} \in \mathcal{N}} \frac{\prod_{j=1}^{2k} \mu(d_j) d_j^{-it_j}}{[d_1, \dots, d_{2k}]} = \prod_{q_{A+1} < p \leq y} \left(1 + \frac{\lambda_p(\mathbf{t})}{p} \right),$$

where

$$(9.13) \quad \lambda_p(\mathbf{t}) := \sum_{I \in S^*(2k)} (-1)^{\#I} p^{-it_I} = -1 + \prod_{j=1}^{2k} (1 - p^{-it_j}).$$

9.3. An auxiliary result. Before we continue with the estimation of $W(\mathbf{J})$, we establish a preliminary (and fairly standard) result.

Lemma 9.1. *For $z \geq y \geq 3$ and $t \in \mathbb{R}$, we have that*

$$\sum_{y < p \leq z} \frac{1}{p^{1+it}} = \begin{cases} \log \left(\frac{\log z}{\log y} \right) + O(1) & \text{if } |t| \leq 1/\log z, \\ \log \left(\frac{1}{|t| \log y} \right) + O(1) & \text{if } 1/\log z < |t| \leq 1/\log y, \\ O(1) & \text{if } y \geq |t| \geq 1/\log y. \end{cases}$$

Finally, if $|t| \geq y$, then

$$\left| \sum_{y < p \leq z} \frac{1}{p^{1+it}} \right| \leq \log \left(\frac{\log(\min\{|t|, z\})}{\log y} \right) + O(1).$$

Proof. If $|t| \leq 1/\log z$, then we note that $p^{-it} = 1 + O(\log p / \log z)$ for $p \leq z$, so that

$$\sum_{y < p \leq z} \frac{1}{p^{1+it}} = \sum_{y < p \leq z} \frac{1}{p} + O(1) = \log \left(\frac{\log z}{\log y} \right) + O(1),$$

as claimed. For $|t| \geq 1/\log y$ and $y \geq |t|$, then we note that

$$\sum_{y < p \leq z} \frac{1}{p^{1+it}} \ll 1$$

by relation (4.4) in [16]. Next, if $1/\log z \leq |t| \leq 1/\log y$, then we apply the results we just proved to deduce that

$$\sum_{y < p \leq e^{1/|t|}} \frac{1}{p^{1+it}} = \log \left(\frac{\log e^{1/|t|}}{\log y} \right) + O(1) = \log \left(\frac{1}{|t| \log y} \right) + O(1)$$

and that

$$\sum_{e^{1/|t|} < p \leq z} \frac{1}{p^{1+it}} \ll 1.$$

Finally, if $|t| \geq y$, then we note that

$$\sum_{|t| < p \leq z} \frac{1}{p^{1+it}} \ll 1,$$

so that

$$\begin{aligned} \left| \sum_{y < p \leq z} \frac{1}{p^{1+it}} \right| &= \left| \sum_{y < p \leq \min\{|t|, z\}} \frac{1}{p^{1+it}} \right| + O(1) \leq \sum_{y < p \leq \min\{|t|, z\}} \frac{1}{p} + O(1) \\ &\leq \log \left(\frac{\log(\min\{|t|, z\})}{\log y} \right) + O(1). \quad \square \end{aligned}$$

9.4. Lower bound for the main term. We now return to the study of the quantity $W(\mathbf{J})$, defined by (9.12). First, we show that the term when $\#J_\ell = 2$ for all ℓ contributes what we claim to be our main term. In this case $L = \#\mathcal{L} = k$ and, by possibly permuting the t_j 's, we may assume that $J_\ell = \{\ell, k + \ell\}$ for all $\ell \in \{1, \dots, k\}$. Thus for these terms we have $t_{k+\ell} = -t_\ell$. We want to show that the integrand in (9.12) is non-negative for all choices of t_1, \dots, t_k . We have that

$$\prod_{j=1}^{2k} \widehat{w}_R(it_j) = \prod_{j=1}^k \widehat{w}_R(it_j) \widehat{w}_R(-it_j) = \prod_{j=1}^k |\widehat{w}_R(it_j)|^2 = \prod_{j=1}^k |t_j^{A+1} \widehat{g}_R(it_j)|^2 \prod_{a=1}^{A+1} \frac{|1 - q_a^{-it_j}|^2}{t_j^2},$$

which is clearly non-negative for all t_1, \dots, t_k . Moreover, if $t_j \ll 1$ for all j , then relations (9.7) and (6.6) imply that

$$\prod_{j=1}^{2k} \widehat{w}_R(it_j) \geq \frac{c_2}{(\log R)^{2kA}}$$

for some $c_2 > 0$. Furthermore, the definition (9.13) implies that in this case we have

$$\lambda_p(\mathbf{t}) = -1 + \prod_{j=1}^k |1 - p^{-it_j}|^2 \geq -1,$$

so

$$\sum_{d_1, \dots, d_{2k} \in \mathcal{N}} \frac{\prod_{j=1}^{2k} \mu(d_j) d_j^{-it_j}}{[d_1, \dots, d_{2k}]} = \prod_{q_{A+1} < p \leq y} \left(1 + \frac{\lambda_p(\mathbf{t})}{p} \right) \geq 0$$

for such \mathbf{t} .

Since the integrand in (9.12) is non-negative, we may obtain a lower bound by restricting the range of integration to any region we wish. We restrict to $t_1 \in [1, 2]$ and $t_j \in [t_1, t_1 + 1/\log y]$ for $2 \leq j \leq k$. The volume of this region is $\asymp 1/(\log y)^{k-1}$. Moreover, in this region we find that

$$\lambda_p(\mathbf{t}) = -1 + |1 - p^{it_1}|^{2k} + O\left(\frac{\log p}{\log y}\right).$$

Therefore

$$\begin{aligned} \sum_{d_1, \dots, d_{2k} \in \mathcal{N}} \frac{\prod_{j=1}^{2k} \mu(d_j) d_j^{-it_j}}{[d_1, \dots, d_{2k}]} &= \exp \left\{ \sum_{p \leq y} \frac{-1 + |1 - p^{it_1}|^{2k}}{p} + O\left(\frac{1}{p^2}\right) + O\left(\frac{\log p}{p \log y}\right) \right\} \\ &\gg \frac{1}{\log y} \exp \left\{ \sum_{p \leq y} \frac{|1 - p^{it_1}|^{2k}}{p} \right\}. \end{aligned}$$

By binomial expansion

$$|1 - p^{it_1}|^{2k} = \sum_{j+j'=2k} \binom{2k}{j} (-1)^{j'} p^{i(j-j')t_1}.$$

The terms with $j = j'$ contribute a factor

$$\exp \left(\sum_{p \leq y} \binom{2k}{k} \frac{1}{p} \right) \asymp (\log y)^{\binom{2k}{k}}.$$

By the final part of Lemma 9.1, since $t_1 \in [1, 2]$ the terms with $j \neq j'$ contribute a factor

$$\exp\left(O\left(\sup_{|j| \leq 2k} \left|\sum_{p \leq y} \frac{p^{ijt_1}}{p}\right|\right)\right) \asymp 1.$$

Thus we obtain a lower bound of $\gg (\log y)^{\binom{2k}{k}-1}$ for our sum over d_1, \dots, d_{2k} . Since the region of integration has volume $\gg (\log y)^{-(k-1)}$, this gives

$$W(\mathbf{J}) \gg \frac{(\log y)^{\binom{2k}{k}-k}}{(\log R)^{k+2kA}}$$

in this case. So we find that the total contribution to the right hand side of (9.10) from such \mathbf{J} is $\geq c_2 (\log y)^{\binom{2k}{k}-k+1} / (\log R)^{(2A+1)k}$ for some $c_2 > 0$, which is greater than the claimed main term in (9.6) if ϵ' is small enough in terms of ϵ and k .

9.5. Upper bound for the error term. It remains to show that the contribution of the \mathbf{J} 's for which at least one of the $\#J_a$'s is odd, is smaller than what we have above. Before we get started, we note that

$$\lambda_p(\mathbf{t}) = -1 + \prod_{j=1}^{2k} (1 - p^{-it_j}) = -1 + \prod_{j=1}^{2k} (p^{it_j/2} - p^{-it_j/2}) = -1 + (-4)^k \prod_{j=1}^{2k} \sin\left(\frac{t_j \log p}{2}\right)$$

whenever $t_1 + \dots + t_{2k} = 0$, which is the case here. In particular, $-4^k \leq \lambda_p(\mathbf{t}) + 1 \leq 4^k$, whence

$$\left| \sum_{d_1, \dots, d_{2k} \in \mathcal{N}} \frac{\prod_{j=1}^{2k} \mu(d_j) d_j^{-it_j}}{[d_1, \dots, d_{2k}]} \right| = \left| \prod_{q_{A+1} < p \leq y} \left(1 + \frac{\lambda_p(\mathbf{t})}{p}\right) \right| \ll \prod_{4^k < p \leq y} \left(1 + \frac{\lambda_p(\mathbf{t})}{p}\right).$$

Next, we split the region of integration in (9.12) into various subsets. First, we note that, by a dyadic decomposition argument and (9.9), we have that

$$W(\mathbf{J}) \ll \frac{(\log T)^{2k}}{(\log R)^{L+2kA}} \max_{1 \leq T_1, \dots, T_{2k} \leq T} \frac{\Lambda(\mathbf{T})}{T_1 \cdots T_{2k-L}} + \frac{(\log R)^{O(1)}}{T},$$

where

$$\Lambda(\mathbf{T}) := \int \cdots \int \prod_{\substack{T_j \leq |t_j|+1 \leq 2T_j \\ 1 \leq j \leq 2k-L \\ t_{J_\ell} = 0 \ (\ell \in \mathcal{L})}} \left(1 + \frac{\lambda_p(\mathbf{t})}{p}\right) dt_1 \cdots dt_{2k-L}.$$

We take

$$T = \exp\{(\log \log R)^2\},$$

and fix a choice of T_1, \dots, T_{2k-L} as above. We will further break the region of integration according to which sums $t_J = \sum_{j \in J} t_j$ are small. Indeed, by Lemma 9.1 the product $\prod_{q_{A+1} < p \leq y} (1 + \lambda_p(\mathbf{t})/p)$ can become large only if there are such configurations. Note that if t_{J_1} and t_{J_2} are both small, so is any linear combination of t_{J_1} and t_{J_2} . Thus, we are naturally led to the following definition: given free variables x_1, \dots, x_{2k} and $J \subset [2k]$, we define the linear form

$$L_J(x_1, \dots, x_{2k}) := \sum_{j \in J} x_j,$$

(thinking of the linear form as acting on the space \mathbb{Q}^{2k}) and, given $\mathcal{I} \subset \mathcal{S}(2k)$, we set

$$V_B(\mathcal{I}) := \left\{ \sum_{I \in \mathcal{I}} \frac{a_I}{q_I} \cdot L_I : a_I, q_I \in \mathbb{Z} \cap [-B, B] \right\}$$

(thought of as a subspace in the dual of \mathbb{Q}^{2k} .) For the simplicity of notation, we also set

$$V(\mathcal{I}) := V_\infty(\mathcal{I}) = \text{Span}_{\mathbb{Q}}(\{L_I : I \in \mathcal{I}\}).$$

Since there are only finitely many linear forms L_I , $I \subset [2k]$, there is some finite $B_0 = B_0(k)$ such that, for any \mathcal{I} , if $L_J \in V(\mathcal{I})$, then $L_J \in V_B(\mathcal{I})$ for all $B \geq B_0$. We assume that $B \geq B_0$ from now on (we will eventually choose B to be large enough in terms of k).

We set $m = \lfloor \log \log y \rfloor$ and, to each \mathbf{t} , we associate the sets

$$\mathcal{I}_j = \mathcal{I}_j(\mathbf{t}) := \{I \subset [2k] : e^m |t_I| + 1 \leq e^{j+1}\} \quad (0 \leq j \leq m).$$

Note that if $L_I \notin V_B(\mathcal{I}_m)$, then $|t_I| > 1$. Since $t_I \ll T$ for all I , Lemma 9.1 implies that

$$\begin{aligned} \prod_{4^k < p \leq y} \left(1 + \frac{\lambda_p(\mathbf{t})}{p}\right) &= \exp\left(\sum_{I \in \mathcal{S}^*(2k)} (-1)^{\#I} \sum_{4^k < p \leq y} \left(p^{-1-it_I} + O\left(\frac{1}{p^2}\right)\right)\right) \\ &= \exp\left(\sum_{\substack{I \in \mathcal{S}^*(2k) \\ |t_I| \leq 1}} (-1)^{\#I} \min\left\{\log \log y, \log\left(\frac{1}{|t_I|}\right)\right\}\right) (\log \log R)^{O(1)} \\ &\ll (\log \log R)^{O(1)} \prod_{\substack{I \in \mathcal{S}^*(2k) \\ L_I \in V_B(\mathcal{I}_m)}} \min\left\{\log y, \frac{1}{|t_I|}\right\}^{(-1)^{\#I}}. \end{aligned}$$

If $L_I \in V_B(\mathcal{I}_j) \setminus V_B(\mathcal{I}_{j-1})$, where $\mathcal{I}_{-1} = \{\emptyset\}$ so that $V_B(\mathcal{I}_{-1}) = \{0\}$, then $I \notin \mathcal{I}_{j-1}$, which means that $e^m |t_I| + 1 > e^j$. On the other hand, since $L_I \in V_B(\mathcal{I}_j)$, then we find that $e^m |t_I| \leq B \cdot \#\mathcal{I}_j \cdot e^{j-m}$. We thus conclude that

$$\min\left\{\log y, \frac{1}{|t_I|}\right\} \asymp e^{m-j} \quad \text{for } L_I \in V_B(\mathcal{I}_j) \setminus V_B(\mathcal{I}_{j-1}), \quad 0 \leq j \leq m.$$

Therefore

$$\prod_{4^k < p \leq y} \left(1 + \frac{\lambda_p(\mathbf{t})}{p}\right) \ll (\log \log R)^{4^k} \prod_{j=0}^m e^{(m-j)F_j},$$

where

$$F_j := \sum_{\substack{I \subset [2k] \\ L_I \in V_B(\mathcal{I}_j) \setminus V_B(\mathcal{I}_{j-1})}} (-1)^{\#I}.$$

(Here we use the fact that there are only finitely many indices j for which $\mathcal{I}_j \neq \mathcal{I}_{j-1}$.) Summing over the $\ll (\log \log R)^{O(1)}$ choices for $\mathcal{I}_0, \mathcal{I}_1, \dots, \mathcal{I}_m$, we conclude that

$$\Lambda(\mathbf{T}) \ll (\log \log R)^{O(1)} \max_{\mathcal{I}_m \supset \dots \supset \mathcal{I}_0 \supset \{J_1, \dots, J_L\}} \left(\nu(\mathbf{T}, \mathcal{I}) \cdot \prod_{j=0}^m e^{(m-j)F_j} \right),$$

where $\nu(\mathbf{T}, \mathcal{I})$ denotes the $(2k - L)$ -dimensional Lebesgue measure of $(t_1, \dots, t_{2k-L}) \in \mathbb{R}$ such that $T_j \leq |t_j| + 1 \leq 2T_j$ for $j \leq 2k - L$ and

$$\{I \subset [2k] : e^m |t_I| + 1 < e^{j+1}\} = \mathcal{I}_j \quad (0 \leq j \leq m),$$

where the variables $t_{2k-L+1}, \dots, t_{2k}$ are defined via the equations $t_{J_\ell} = 0$ for $\ell \in \{1, \dots, 2k\}$. (In particular, $\mathcal{I}_0 \supset \{J_1, \dots, J_L\}$.)

Next, we use linear algebra to understand $\nu(\mathbf{T}, \mathcal{I})$. If

$$D_j = \dim V(\mathcal{I}_j),$$

then we may find sets I_1, \dots, I_{D_m} such that, for each $j \leq m$, $\{L_{I_1}, \dots, L_{I_{D_j}}\}$ is a basis of $V(\mathcal{I}_j)$. We may also assume that $\{J_\ell : 1 \leq \ell \leq L\}$ is contained in $\{I_1, \dots, I_{D_1}\}$.

Eliminating variables from linear combinations of the asymptotic formulas

$$t_I = O(e^{j-m}) \quad (I \in \{I_1, \dots, I_{D_j}\})$$

(for example, as in Gaussian elimination), yields D_j of the variables t_i , say the variables $\{t_i : i \in \mathcal{D}_j\}$ (where $\#\mathcal{D}_j = D_j$), in terms of linear combinations of the other variables, up to an error of $O(e^{j-m})$. We can also arrange the sets $\mathcal{D}_0, \dots, \mathcal{D}_L$ to satisfy $\mathcal{D}_0 \subset \dots \subset \mathcal{D}_L$. (Recall that $\mathcal{I}_j \neq \mathcal{I}_{j-1}$ for only finitely many j 's.) We may therefore prove that

$$\nu(\mathbf{T}, \mathbf{I}) \ll (\log y)^L \left(\prod_{j=0}^m e^{(D_j - D_{j-1})(j-m)} \right) \left(\prod_{\substack{1 \leq j \leq 2k \\ j \notin \mathcal{D}_0 \cup \dots \cup \mathcal{D}_L}} T_j \right),$$

where the extra factor $(\log y)^L$ is included because we are not integrating over the variables $t_{2k-L+1}, \dots, t_{2k}$, which are fixed via the conditions $t_{J_\ell} = 0$, $1 \leq \ell \leq L$. By the above discussion, we find that

$$\Lambda(\mathbf{T}) \ll T_1 \cdots T_{2k-L} (\log y)^L (\log \log R)^{O(1)} \max_{\mathcal{I}_m \supset \dots \supset \mathcal{I}_0 \supset \{J_1, \dots, J_L\}} \prod_{j=0}^m e^{(m-j)(F_j - (D_j - D_{j-1}))}.$$

We note that

$$\begin{aligned} \sum_{j=0}^m (m-j)(F_j - (D_j - D_{j-1})) &= \sum_{j=0}^{m-1} \sum_{i=1}^{m-j} (F_j - (D_j - D_{j-1})) \\ &= \sum_{i=1}^m \sum_{j=0}^{m-i} (F_j - (D_j - D_{j-1})) \\ &= \sum_{i=1}^m (\mathcal{A}(V(\mathcal{I}_{m-i})) - \dim(V(\mathcal{I}_{m-i}))) \end{aligned}$$

in the notation of Section 7, provided that B is large enough. Since \exp is a convex function, we have that

$$e^{\sum_{j=0}^m (m-j)(F_j - (D_j - D_{j-1}))} \leq \frac{1}{m} \sum_{j=0}^{m-1} e^{m(\mathcal{A}(V(\mathcal{I}_j)) - \dim(V(\mathcal{I}_j)))} \asymp \frac{1}{m} \sum_{j=0}^{m-1} (\log y)^{\mathcal{A}(V(\mathcal{I}_j)) - \dim(V(\mathcal{I}_j))}$$

We then conclude that

$$W(\mathbf{J}) \ll \frac{(\log \log R)^{O(1)} (\log y)^L}{(\log R)^{L+2kA}} \max_{\mathcal{I} \supset \{J_1, \dots, J_L\}} (\log y)^{\mathcal{A}(V(\mathcal{I})) - \dim(V(\mathcal{I}))}.$$

The above discussion reduces (9.6) to proving that

$$(9.14) \quad \mathcal{A}(V(\mathcal{I})) - \dim(V(\mathcal{I})) \leq \binom{2k}{k} - 2k - 1$$

whenever \mathcal{I} contains the sets J_1, \dots, J_L and at least one of the J_i 's has an odd number of elements. This follows directly by Proposition 7.1, thus completing the proof of (9.1) and, hence, of Theorem 1.3.

10. ON THE ANATOMY OF INTEGERS CONTRIBUTING TO $\mathcal{M}_{f,2k}(R)$

This section is devoted to establishing Theorems 1.6 and 1.4. Throughout this section, given $n \in \mathbb{Z}_{\geq 1}$ and $R \geq 1$, we adopt the notations

$$\Omega(n; R) := \sum_{p^\alpha \parallel n, p \leq R} \alpha \quad \text{and} \quad \Omega(n; r, R) := \sum_{p^\alpha \parallel n, r < p \leq R} \alpha.$$

A key observation, that we will use several times, is that if $(a, b) = 1$, then

$$\begin{aligned} M_{f_A}(ab; R) &= \sum_{d|a} \sum_{d'|b} \mu(d) \mu(d') f_A \left(\frac{\log d}{\log R} + \frac{\log d'}{\log R} \right) \\ (10.1) \quad &= \sum_{\substack{d'|b \\ d \leq R}} \mu(d') \left(\frac{\log(R/d')}{\log R} \right)^A \sum_{d|a} \mu(d) f_A \left(\frac{\log d}{\log(R/d')} \right) \end{aligned}$$

by the definition of f_A .

10.1. An estimate for the logarithmic means. We start by proving a preliminary result, where the integer n is weighted by $1/n$. The transition to the uniform weights will be accomplished in the subsequent section.

Theorem 10.1. *Let $R \geq 2$, $k \in \mathbb{Z}_{\geq 1}$ and $A \in \mathbb{Z}_{\geq 0}$.*

(a) *If $A > \frac{1}{2k} \binom{2k}{k} - 1$ and $\eta \in [\log 2 / \log R, 1]$, then*

$$\prod_{p \leq R} \left(1 - \frac{1}{p} \right) \sum_{P^+(n) \leq R} \frac{\Omega(n; R^\eta) M_{f_A}(n; R)^{2k}}{n} \ll_{k,A} \frac{\eta^{2k}}{\log R} + (\log R)^{\binom{2k}{k} - 2k(A+1)} \log \log R.$$

(b) *If $A \leq \frac{1}{2k} \binom{2k}{k} - 1$ and $1 - 1/\binom{2k}{k} \leq v \leq 2 - \epsilon$, then*

$$\prod_{p \leq R} \left(1 - \frac{1}{p} \right) \sum_{P^+(n) \leq R} \frac{v^{\Omega(n)} M_{f_A}(n; R)^{2k}}{n} \ll_{\epsilon,k,A} (\log R)^{v \binom{2k}{k} - 2k(A+1)} (\log \log R)^{O_k(1)}.$$

Proof. To ease notation, for this proof we let all implied constants depend on k, A and ϵ without explicitly stating this.

(a) First of all, we claim we may restrict our attention to small enough η . To see this, it suffices to show that

$$(10.2) \quad \sum_{P^+(n) \leq R} \frac{\Omega(n; R^\delta, R) M_{f_A}(n; R)^{2k}}{n} \ll_\delta 1,$$

for any fixed $\delta > 0$. Now, Hölder's inequality applied to (10.1) yields

$$(10.3) \quad M_{f_A}(ab; R)^{2k} \leq \tau(b)^{2k-1} \sum_{\substack{d'|b \\ d' \leq R}} \mu^2(d') \left(\frac{\log(R/d')}{\log R} \right)^{2kA} M_{f_A}(a; R/d')^{2k}.$$

Therefore, writing $n = ab$, where $b = \prod_{p^v \parallel n, R^\delta < p \leq R} p^v$, we find that

$$\begin{aligned} & \sum_{P^+(n) \leq R} \frac{\Omega(n; R^\delta, R) M_{f_A}(n; R)^{2k}}{n} \\ & \leq \sum_{p|b \Rightarrow R^\delta < p \leq R} \frac{\tau(b)^{2k-1} \Omega(b)}{b} \sum_{\substack{d'|b \\ d' \leq R}} \mu^2(d') \left(\frac{\log(R/d')}{\log R} \right)^{2kA} \sum_{P^+(a) \leq R} \frac{M_{f_A}(a; R/d')^{2k}}{a} \\ & \ll \sum_{p|b \Rightarrow R^\delta < p \leq R} \frac{\tau(b)^{2k-1} \Omega(b)}{b} \sum_{\substack{d'|b \\ d' \leq R}} \mu^2(d') \left(\frac{\log(R/d')}{\log R} \right)^{2kA} \frac{\log R}{\log(R/d')} \ll_\delta 1 \end{aligned}$$

by Theorem 1.3, since $A > \frac{1}{2k} \binom{2k}{k} - 1$ and so $\mathcal{E}_{k,A} = -1$ and $A \geq 1$. This proves (10.2), so for the rest of the proof, we may assume that η is sufficiently small.

Expanding $\Omega(n, R^\eta)$, we have that

$$\begin{aligned} S &:= \prod_{p \leq R} \left(1 - \frac{1}{p} \right) \sum_{P^+(n) \leq R} \frac{\Omega(n; R^\eta) M_{f_A}(n; R)^{2k}}{n} \\ &= \prod_{p \leq R} \left(1 - \frac{1}{p} \right) \sum_{\substack{q \leq R^\eta \\ j \geq 1}} \frac{j}{q^j} \sum_{\substack{P^+(m) \leq R \\ q \nmid m}} \frac{M_{f_A}(mq^j; R)^{2k}}{m} \\ &\leq \prod_{p \leq R} \left(1 - \frac{1}{p} \right) \sum_{q \leq R^\eta} \frac{q}{(q-1)^2} \sum_{P^+(m) \leq R} \frac{\left(\sum_{d|m} \mu(d) \left\{ f_A \left(\frac{\log d}{\log R} \right) - f_A \left(\frac{\log(qd)}{\log R} \right) \right\} \right)^{2k}}{m}. \end{aligned}$$

where we used (10.1) with $a = q^j$ and $b = m$. Expanding the $2k^{th}$ power, we find that

$$S \leq \sum_{q \leq R^\eta} \frac{q}{(q-1)^2} \sum_{\substack{P^+(d_j) \leq R \\ 1 \leq j \leq 2k}} \frac{\prod_{j=1}^{2k} \mu(d_j) \left\{ f_A \left(\frac{\log d_j}{\log R} \right) - f_A \left(\frac{\log(qd_j)}{\log R} \right) \right\}}{[d_1, \dots, d_{2k}]},$$

Here $A \geq 1$, so that $\widehat{(f_A)}(s) = A! R^s / (s^{A+1} (\log R)^A)$ decays fast, and Perron inversion implies that

(10.4)

$$S \leq \sum_{q \leq R^\eta} \frac{q}{(q-1)^2} \int \cdots \int_{\substack{\operatorname{Re}(s_j) = 1/\log R \\ 1 \leq j \leq 2k}} \sum_{d_1, \dots, d_{2k} \geq 1} \frac{\prod_{j=1}^{2k} \mu(d_j) d_j^{-s_j}}{[d_1, \dots, d_{2k}]} \prod_{j=1}^{2k} \frac{A! R^{s_j} (1 - q^{-s_j})}{(\log R)^A s_j^{A+1}} ds_1 \cdots ds_{2k}.$$

The above integral is amenable to the methods of Section 8. Precisely, we note that the integrand is of the form

$$A!^{2k} \left(\frac{\log q}{\log R} \right)^{2k} \frac{P(\mathbf{s}) R^{s_{[2k]}}}{(\log R)^{2k(A-1)}} \prod_{I \in \mathcal{S}^+(2k)} \zeta^{e_I} (1 + s_I) \prod_{j=1}^{2k} \frac{(1 - q^{-s_j}) / (s_j \log q)}{(s_j \zeta (1 + s_j))^A},$$

where $P(\mathbf{s})$ is as in Section 8, $e_I = +1$ for $I \in \mathcal{S}^+(2k)$, $e_I = -1$ for $I \in \mathcal{S}^-(2k)$ with $\#I > 1$, and $e_I = A - 1 \geq 0$ for $\#I = 1$. If $q \leq R^\delta$ with δ small enough, then the argument leading to

(8.1) yields that

$$(10.5) \quad \int \cdots \int_{\substack{\operatorname{Re}(s_j)=1/\log R \\ 1 \leq j \leq 2k}} \sum_{d_1, \dots, d_{2k} \geq 1} \frac{\prod_{j=1}^{2k} \mu(d_j) d_j^{-s_j}}{[d_1, \dots, d_{2k}]} \prod_{j=1}^{2k} \frac{A! R^{s_j} (1 - q^{-s_j})}{(\log R)^A s_j^{A+1}} ds_1 \cdots ds_{2k} \\ \ll \frac{(\log q)^{2k}}{(\log R)^{2k+1}} + (\log R)^{\binom{2k}{k} - 2k(A+1)},$$

with the first term coming from Case 1a and the second one from Case 2. Inserting the above inequality into (10.4) completes the proof of part (a).

(b) We will use a variation of the argument of Section 9. The fact that Proposition 7.1 requires $s_{[2k]} = 0$ complicates the proof; otherwise, a direct application of the method of Section 9.5 would be possible.

Call S' the sum in question. As usual, we may replace f_0 by a sufficiently smooth function h . So write $g = h$ if $A = 0$, and $g = f_A$ otherwise. Note that, since $M_g(n; R)$ depends only on the square-free kernel of n , we have that

$$S' = \prod_{p \leq R} \left(1 - \frac{1}{p}\right) \sum_{P^+(n) \leq R} \frac{\mu^2(n) v^{\omega(n)}}{\varphi_v(n)} M_{f_A}(n; R)^{2k}$$

with $\varphi_v(n) = \prod_{p|n} (p - v)$. Set $y = R^{c/\log \log R}$ for a small enough but fixed c . Given an integer n , we decompose it as $n = ab$ with $P^+(a) \leq y < P^-(b)$. If $\omega(a) \leq A + 1$, then $M_{f_A}(n; R) \ll 4^{k\omega(b)}$, and we immediately find that such n 's contribute at most $\ll (\log \log R)^{O(1)} / \log R$. Otherwise, we sum over all possible choices $a = qa'$ with $\omega(q) = A + 1$ to deduce that

$$(10.6) \quad S' \leq v^{A+1} \sum_{\substack{P^+(q) \leq y \\ \omega(q) = A+1}} \frac{\mu^2(q)}{\varphi_v(q)} S'(q) + O\left(\frac{(\log \log R)^{O(1)}}{\log R}\right),$$

where

$$S'(q) := \prod_{p \leq R} \left(1 - \frac{1}{p}\right) \sum_{\substack{P^+(a') \leq y < P^-(b) \leq R \\ (a', q) = 1}} \frac{\mu^2(a'b) v^{\omega(a'b)}}{\varphi_v(a'b)} M_{f_A}(qa'b; R)^{2k}.$$

Next, we apply (10.3) with $a = a'q$ to find that

$$S'(q) \leq \prod_{p \leq R} \left(1 - \frac{1}{p}\right) \sum_{\substack{P^+(a') \leq y \\ y < P^-(b) \leq R \\ (a', q) = 1}} \frac{\mu^2(a'b) v^{\omega(a'b)} \tau(b)^{2k-1}}{\varphi_v(a'b)} \sum_{\substack{d|b \\ d \leq R}} \left(\frac{\log(R/d)}{\log R}\right)^{2kA} M_{f_A}(a'q; R/d)^{2k}.$$

We write $b = cd$ and note that the sum over c is $\leq (\log \log R)^{O(1)}$ by our choice of y . Moreover, $\varphi_v(n) \gg n/(\log \log n)^v$. Thus,

$$S'(q) \ll (\log \log R)^{O(1)} \sum_{\substack{d \leq R \\ P^-(d) > y}} \frac{4^{k\Omega(d; y, R)}}{d} S''(d, q) \leq (\log \log R)^{O(1)} \sum_{\substack{d \leq R \\ P^-(d) > y}} \frac{S'''(d, q)}{d^{1-2k/\log y}},$$

where

$$\begin{aligned} S''(d, q) &:= \prod_{p \leq R} \left(1 - \frac{1}{p}\right) \sum_{\substack{P^+(a') \leq y \\ (a', q)=1}} \frac{\mu^2(a') v^{\omega(a')}}{\varphi_v(a')} \left(\frac{\log(R/d)}{\log R}\right)^{2kA} M_{f_A}(a'q; R/d)^{2k} \\ &\leq \prod_{p \leq R} \left(1 - \frac{1}{p}\right) \sum_{\substack{P^+(a) \leq y \\ (a, q)=1}} \frac{v^{\Omega(a)}}{a} \left(\frac{\log(R/d)}{\log R}\right)^{2kA} M_{f_A}(aq; R/d)^{2k}. \end{aligned}$$

Before we proceed to the estimation of $S''(d, q)$, we note that

$$(10.7) \quad S'(q) \ll (\log \log R)^{O(1)} \sum_{d \leq R} \frac{(\lambda^+ * 1)(d)}{d^{1-2k/\log y}} S''(d, q),$$

where $(\lambda^+(m))_{m \leq R^\delta}$ is an upper bound sieve with δ small enough, constructed using the fundamental lemma of sieve methods, taking $\kappa = 1$ in [14, Lemma 6.3, p. 159]. Then the Dirichlet series $\sum_d (1 * \lambda^+)(d)/d^s$ has a simple pole at $s = 1$ of residue $\sum_m \lambda^+(m)/m \asymp (\log \log R)^{O(1)}/\log R$.

Next, if $q = p_1^{r_1} \cdots p_{A+1}^{r_{A+1}}$ is the prime factorisation of q , then (10.1) implies that

$$\left(\frac{\log(R/d)}{\log R}\right)^A M_{f_A}(aq; R/d) = \sum_{d'|aq} \mu(d') f_A\left(\frac{\log(dd')}{\log R}\right) = \sum_{d''|a} \mu(d'') w_q\left(\frac{\log(dd'')}{\log R}\right)$$

with

$$w_q(x) := \sum_{J \subset [A+1]} (-1)^{\#J} f_A\left(x + \frac{\sum_{j \in J} \log p_j}{\log R}\right).$$

As usual, if $A = 0$, we may replace f_0 by a sufficiently smooth function h at the cost of a small error. Letting $g = h$ when $A = 0$ and $g = f_A$ otherwise, and letting W_q have the same definition as w_q with g in place of f_A , we find that

$$\begin{aligned} S'''(d, q) &\leq \prod_{p \leq R} \left(1 - \frac{1}{p}\right) \sum_{\substack{P^+(a) \leq y \\ (a, q)=1}} \frac{v^{\Omega(a)}}{a} \left(\sum_{f|a} \mu(f) W_q\left(\frac{\log(df)}{\log R}\right)\right)^{2k} + O\left(\frac{1}{\log R}\right) \\ &= \frac{\prod_{p \leq R} (1 - 1/p)}{\prod_{p \leq y, p \nmid q} (1 - v/p)} \sum_{\substack{(f_j, q)=1 \\ 1 \leq j \leq 2k}} \frac{v^{\Omega([f_1, \dots, f_{2k}])}}{[f_1, \dots, f_{2k}]} \prod_{j=1}^{2k} \mu(f_j) W_q\left(\frac{\log(df_j)}{\log R}\right) + O\left(\frac{1}{\log R}\right). \end{aligned}$$

We apply Mellin inversion $2k$ times to find that

$$\begin{aligned} S'''(d, q) &\leq \frac{\prod_{p \leq R} (1 - 1/p)}{(2\pi i)^{2k} \prod_{p \leq y, p \nmid q} (1 - v/p)} \int \cdots \int \sum_{\substack{P^+(f_j) \leq y, (f_j, q)=1 \\ \operatorname{Re}(s_j) = 4k/\log y \\ |\operatorname{Im}(s_j)| \leq T \\ 1 \leq j \leq 2k}} \frac{v^{\Omega([f_1, \dots, f_{2k}])} \prod_{j=1}^{2k} \mu(f_j) f_j^{-s_j}}{[f_1, \dots, f_{2k}]} \\ &\quad \times \prod_{j=1}^{2k} \frac{\widehat{g}_R(s_j)}{d^{s_j}} \prod_{a=1}^{A+1} (1 - p_a^{-s_j}) ds_j + O\left(\frac{1}{\log R}\right). \end{aligned}$$

Together with (10.7), this implies that

$$S'(q) \leq \frac{(\log \log R)^{O(1)} \prod_{p \leq R} (1 - 1/p)}{(2\pi i)^{2k} \prod_{p \leq y, p \nmid q} (1 - v/p)} \int \cdots \int_{\substack{\operatorname{Re}(s_j) = 4k/\log y \\ |\operatorname{Im}(s_j)| \leq T \\ 1 \leq j \leq 2k}} \sum_{\substack{P^+(f_j) \leq y, (f_j, q) = 1 \\ 1 \leq j \leq 2k}} \frac{v^{\Omega([f_1, \dots, f_{2k}])} \prod_{j=1}^{2k} \mu(f_j) f_j^{-s_j}}{[f_1, \dots, f_{2k}]} \\ \times P(1 + s_{[2k]} - 2k/\log y) \prod_{j=1}^{2k} \widehat{g}_R(s_j) \prod_{a=1}^{A+1} (1 - p_a^{-s_j}) ds_j + O\left(\frac{1}{\log R}\right),$$

where

$$P(s) := \sum_{d=1}^{\infty} \frac{(\lambda^+ * 1)(d)}{d^s} = \zeta(s) \sum_{m \leq R^\delta} \frac{\lambda^+(m)}{m^s}.$$

We fix s_1, \dots, s_{2k-1} and move s_{2k} to the line $\operatorname{Re}(s_{2k}) = -8k/\log y$ to pick up the pole at $s_{[2k]} = 2k/\log y$. If c is small enough in the definition of y , and the same is true for δ , the complementary contours contribute $\ll 1/\log R$. There are no other poles, since the factors $1 - p_a^{-s_j}$ are annihilating the poles of $\widehat{g}_R(s_j)$. We conclude that

$$S'(q) \ll \frac{(\log \log R)^{O(1)}}{(\log R)^{2kA+2-v}} \int \cdots \int_{\substack{\operatorname{Re}(s_j) = 4k/\log y \\ |\operatorname{Im}(s_j)| \leq T \\ 1 \leq j \leq 2k \\ s_{[2k]} = 2k/\log y}} \left| \sum_{\substack{P^+(f_j) \leq y \\ 1 \leq j \leq 2k}} \frac{v^{\Omega([f_1, \dots, f_{2k}])} \prod_{j=1}^{2k} \mu(f_j) f_j^{-s_j}}{[f_1, \dots, f_{2k}]} \right| \\ \left(\prod_{j=1}^{2k} \frac{1}{|s_j|^{A+1}} \right) |ds_1 \cdots ds_{2k-1}| + \frac{1}{\log R}.$$

Recall the notation $\lambda_p(\mathbf{t})$ defined in (9.13), and combine the above with (10.6) to find that

$$S' \ll \frac{(\log \log R)^{O(1)}}{(\log R)^{2kA+2-v}} \int \cdots \int_{\substack{-T \leq t_j \leq T \ (1 \leq j < 2k) \\ t_{[2k]} = 0}} \left(1 + \frac{v \cdot \lambda_p(\mathbf{t}) + (A+1) \sum_{j=1}^{2k} \cos(t_j \log p)}{p} \right) \\ \times \left(\prod_{j=1}^{2k} \frac{(\log(2 + |t_j|))^{2k(A+1)}}{(1 + |t_j|)^{A+1}} \right) dt_1 \cdots dt_{2k-1} + \frac{(\log \log R)^{O(1)}}{\log R},$$

where we used the fact that $s\zeta(1+s) \gg 1/\log(2+|t|)$ for $\operatorname{Re}(s) > 1$. Then, following the arguments of Section 9.5 (with $L = 1$), and recalling the notations $V(\mathcal{I})$ and $\mathcal{A}(V(\mathcal{I}))$, we find that

$$S' \ll \max_{\mathcal{I} \subset \mathcal{I}(2k)} (\log R)^{v(1+\mathcal{A}(V(\mathcal{I}))) - \dim(V(\mathcal{I})) + (A+1)U(\mathcal{I}) - 2kA - 1} (\log \log R)^{O(1)},$$

where

$$U(\mathcal{I}) := \#\{1 \leq j \leq 2k : L_{\{j\}} \in V(\mathcal{I})\}.$$

If $U(\mathcal{I}) \geq 1$, then Proposition 7.1(a) implies that $\mathcal{A}(V(\mathcal{I})) = -1$, and the exponent of $\log R$ is then

$$(A+1) \cdot U(\mathcal{I}) - \dim(V(\mathcal{I})) - 2kA - 1 \leq A \cdot U(\mathcal{I}) - 2kA - 1 \leq -1,$$

since we clearly have that $2k \geq \dim(V(\mathcal{I})) \geq U(\mathcal{I})$. Assume now that $U(\mathcal{I}) = 0$. If $\mathcal{A}(V(\mathcal{I})) = \binom{2k}{k} - 1$ and $\dim(V(\mathcal{I})) = 2k - 1$, then the exponent of $\log R$ is $v\binom{2k}{k} - 2k(A+1)$. Finally, if this

is not the case, then Proposition 7.1 (together with the argument in the end of Section 9.5) implies that

$$\begin{aligned} v(1 + \mathcal{A}(V(\mathcal{I}))) - \dim(V(\mathcal{I})) - 1 &\leq \max\{0, v - 1\} \binom{2k}{k} + \mathcal{A}(V(\mathcal{I})) - \dim(V(\mathcal{I})) \\ &\leq \max\{0, v - 1\} \binom{2k}{k} + \binom{2k}{k} - 2k - 1 \\ &\leq v \binom{2k}{k} - 2k, \end{aligned}$$

by our assumption that $v \geq 1 - 1/\binom{2k}{k}$. This completes the proof of the theorem. \square

10.2. From logarithmic weights to uniform weights. In this section, we show how to go from Theorem 10.1 to the analogous result for the regular mean value and then prove Theorem 1.6.

Theorem 10.2. *Let $x \geq R \geq 2$, $k \in \mathbb{Z}_{\geq 1}$, $A \in \mathbb{Z}_{\geq 0}$ and $\epsilon \in (0, 1/2)$.*

(a) *Assume that $A > \frac{1}{2k} \binom{2k}{k} - 1$. Then uniformly for $\eta \in [\log 2 / \log R, 1]$, we have*

$$\frac{1}{x} \sum_{n \leq x} \Omega(n; R^\eta) M_{f_A}(n; R)^{2k} \ll_{k,A} \frac{\eta}{\log R}.$$

(b) *If $A \leq \frac{1}{2k} \binom{2k}{k} - 1$ and $1 - 1/\binom{2k}{k} + \epsilon \cdot \mathbf{1}_{k=1} \leq v \leq 2 - \epsilon$, then*

$$\frac{1}{x} \sum_{n \leq x} v^{\Omega(n; R)} M_{f_A}(n; R)^{2k} \ll_{k,A,\epsilon} (\log R)^{v \binom{2k}{k} - 2k(A+1)} (\log \log R)^{O(1)}.$$

Proof. We start by proving a preparatory estimate. Our goal is to show that there is some constant $C = C(k)$ such that

$$(10.8) \quad \sum_{n \leq x} M_{f_A}(n; R)^{2k} \leq \frac{Cx}{\log R} \quad (x \geq R > 1).$$

When $x \leq 2^H$ for some fixed $H \in \mathbb{Z}_{\geq 1}$ that will be taken large enough in terms of k and A , relation (10.8) trivially holds by taking C to be large enough in terms of H . Assume now that (10.8) holds for all $x \leq 2^h$, where $h \geq H$. We wish to prove that (10.8) is also true when $x \in (2^h, 2^{h+1}]$.

We follow a variation of the argument leading to Theorem III.3.5 in [25, p.308]: note that

$$(10.9) \quad \sum_{n \leq x} M_{f_A}(n; R)^{2k} = \sum_{n \leq x} M_{f_A}(n; R)^{2k} \frac{\log(x/n)}{\log x} + \sum_{n \leq x} M_{f_A}(n; R)^{2k} \frac{\log n}{\log x}.$$

For the first sum, we bound $\log(x/n)$ by x/n to give

$$\begin{aligned} \sum_{n \leq x} M_{f_A}(n; R)^{2k} \frac{\log(x/n)}{\log x} &\leq \sum_{n \leq x} M_{f_A}(n; R)^{2k} \frac{x}{n \log x} \\ (10.10) \quad &\ll \frac{x}{\log R} \sum_{P^+(n) \leq R} \frac{M_{f_A}(n; R)^{2k}}{n} \\ &\ll \frac{x}{\log R}, \end{aligned}$$

by Theorem 1.3. In the second sum, we write $\log n = \sum_{p^j \parallel n} j \log p$ to find that

$$\sum_{n \leq x} M_{f_A}(n; R)^{2k} \frac{\log n}{\log x} = \sum_{p^j \leq x} \frac{j \log p}{\log x} \sum_{\substack{m \leq x/p^j \\ p \nmid m}} M_{f_A}(mp^j; R)^{2k}.$$

Since

$$M_{f_A}(mp^j; R) = M_{f_A}(m; R) - \mathbf{1}_{p < R} \left(\frac{\log(R/p)}{\log R} \right)^A M_{f_A}(m; R/p)$$

by (10.1), Minkowski's inequality implies that

$$\sum_{n \leq x} M_{f_A}(n; R)^{2k} \frac{\log n}{\log x} \leq \left(S_1^{\frac{1}{2k}} + S_2^{\frac{1}{2k}} \right)^{2k},$$

where

$$S_1 := \sum_{p^j \leq x} \frac{j \log p}{\log x} \sum_{m \leq x/p^j} M_{f_A}(m; R)^{2k}.$$

and

$$S_2 := \sum_{\substack{p^j \leq x \\ p < R}} \frac{j \log p}{\log x} \left(\frac{\log(R/p)}{\log R} \right)^{2kA} \sum_{m \leq x/p^j} M_{f_A}(m; R/p)^{2k}.$$

For S_1 , we note that

$$S_1 = \sum_{m \leq x} M_{f_A}(m; R)^{2k} \sum_{p^j \leq x/m} \frac{j \log p}{\log x} \ll \frac{x}{\log x} \sum_{m \leq x} \frac{M_{f_A}(m; R)^{2k}}{m} \ll \frac{x}{\log R},$$

by (10.10). Finally, we need to bound S_2 . First, we bound its subsum with $j \geq 2$. We have that

$$\begin{aligned} & \sum_{\substack{p^j \leq x \\ j \geq 2 \\ p < R}} \frac{j \log p}{\log x} \left(\frac{\log(R/p)}{\log R} \right)^{2kA} \sum_{m \leq x/p^j} M_{f_A}(m; R/p)^{2k} \\ & \leq \sum_{\substack{p^j \leq x \\ j \geq 2 \\ p < R}} \frac{j \log p}{\log x} \left(\frac{\log(R/p)}{\log R} \right)^{2kA} \sum_{m \leq x/p^j} M_{f_A}(m; R/p)^{2k} \frac{x}{p^j m} \\ & \ll x \sum_{\substack{p^j \leq x \\ j \geq 2 \\ p < R}} \frac{j \log p}{p^j \log x} \left(\frac{\log(R/p)}{\log R} \right)^{2kA} \cdot \frac{\log x}{\log(R/p)}, \end{aligned}$$

by (10.10) with R replaced by R/p . Since $A \geq 1$ here, we find that the above is

$$\ll \sum_{\substack{p^j \leq x \\ j \geq 2}} \frac{j \log p}{p^j} \cdot \frac{x}{\log R} \ll \frac{x}{\log R},$$

where the implied constant is independent of C .

Finally, we need to bound the part of S_2 with $j = 1$. We note that $x/p \leq 2^h$ and that $R/p \leq x/p$, so we may apply the induction hypothesis. This gives a bound

$$\begin{aligned} &\leq \sum_{p < R} \frac{\log p}{\log x} \left(\frac{\log(R/p)}{\log R} \right)^{2kA} \cdot \frac{Cx}{p \log(R/p)} \leq \frac{Cx}{(\log x)(\log R)^2} \sum_{p < R} \frac{(\log p)(\log(R/p))}{p} \\ &\leq \frac{2C}{3} \cdot \frac{x}{\log R}, \end{aligned}$$

provided that H (and hence x) is big enough, where we used our assumptions that $A \geq 1$ and $x \geq R$. Combining the above, and assuming that C is big enough in terms of k and A completes the inductive step and hence the proof of (10.8).

We now turn to proving part (a), that is bounding the sum

$$S(x, R, Q) := \sum_{n \leq x} \Omega(n; Q) M_{f_A}(n; R)^{2k}.$$

The proof is similar to the proof of (10.8), so we only give the main technical twists: we use induction on the dyadic interval in which x lies to prove that there is some constant $C' = C'(k, \epsilon)$ such that

$$(10.11) \quad S(x, R, Q) \leq \frac{C' x \log Q}{(\log R)^2} \quad (x \geq R \geq Q \geq 2).$$

When $x \leq 2^H$, for some fixed $H \in \mathbb{Z}_{\geq 1}$ that will be taken large enough in terms of k, A and ϵ , this trivially holds by taking C' to be large enough in terms of H . Assume now that (10.8) holds for all $x \leq 2^h$, where $h \geq H$. We will prove that it also holds for $x \in (2^h, 2^{h+1}]$. Note that the analogues of (10.9) and (10.10) hold here as well, so let us focus on understanding the sum

$$T := \sum_{n \leq x} \Omega(n; Q) M_{f_A}(n; R)^{2k} \frac{\log n}{\log x}.$$

Fix a large integer N and call T_1 the portion of this sum with $\Omega(n; Q) > 2N$ and T_2 the remaining sum. Writing $\log n = \sum_{p^j \parallel n} j \log p$, we find that

$$T_1 = \sum_{p^j \leq x} \frac{j \log p}{\log x} \sum_{\substack{m \leq x/p^j, p \nmid m \\ \Omega(mp^j; Q) > 2N}} \Omega(mp^j; Q) M_{f_A}(mp^j; R)^{2k}.$$

If T'_1 is the part with $\Omega(mp^j; Q) \leq 2j$ and T''_1 is the rest, then

$$\begin{aligned}
T'_1 &\leq \sum_{\substack{p^j \leq x \\ j > N}} \frac{2j^2 \log p}{\log x} \sum_{\substack{m \leq x/p^j \\ p \nmid m}} M_{f_A}(mp^j; R)^{2k} \frac{x}{p^j m} \\
&\ll x \sum_{\substack{p^j \leq x, p \leq e^{\sqrt{\log R}} \\ j > N}} \frac{j^2 \log p}{p^j \log R} \sum_{\substack{P^+(m) \leq R \\ p \nmid m}} \frac{M_{f_A}(mp; R)^{2k}}{m} + \frac{x(\log R)^{O(1)}}{e^{\sqrt{\log R}}} \\
&\ll_N x \sum_{\substack{p^j \leq x, p \leq e^{\sqrt{\log R}} \\ j > N}} \frac{j^2 \log p}{p^j \log R} \cdot \left\{ \left(\frac{\log p}{\log R} \right)^{2k} + (\log R)^{1 + \binom{2k}{k} - 2k(A+1)} \right\} \\
&\ll \frac{x}{(\log R)^{2k+1}} + x(\log R)^{\binom{2k}{k} - 2k(A+1)},
\end{aligned}$$

where the second to last bound follows from (10.5). Finally, in the range of T''_1 , we note that $\Omega(m; Q) \geq (\Omega(m; Q) + j)/2 > N$, so that $\Omega(mp^j; Q) \leq j(1 + \Omega(m; Q)) \leq (1 + 1/N) \cdot j \cdot \Omega(m; Q)$. Therefore,

$$\begin{aligned}
(10.12) \quad T_1 &\leq \frac{N+1}{N} \sum_{p^j \leq x} \frac{j^2 \log p}{\log x} \sum_{\substack{m \leq x/p^j, p \nmid m \\ \Omega(mp^j; Q) > 2N}} \Omega(m; Q) M_{f_A}(mp^j; R)^{2k} \\
&\quad + O\left(\frac{x}{(\log R)^{2k+1}} + x(\log R)^{\binom{2k}{k} - 2k(A+1)} \right).
\end{aligned}$$

Next, we need to bound

$$T_2 = \sum_{\substack{n \leq x \\ \Omega(n; Q) \leq 2N}} \Omega(n; Q) M_{f_A}(mp^j; R)^{2k} \frac{\log n}{\log x}.$$

If $Q > R^{1/(2N^2)}$, then we simply note that

$$T_2 \leq 2N \sum_{n \leq x} M_{f_A}(n; R)^{2k} \ll_N \frac{x}{\log R} \ll_N \frac{x \log Q}{(\log R)^2}.$$

by (10.8). Here the implied constant depends on N but does not depend on C' .

On the other hand, if $Q \leq R^{1/(2N^2)}$, then we have that $\sum_{p^j \parallel n, p \leq Q} j \log p \leq (\log x)/N$, so that

$$\begin{aligned}
T_2 &\leq \frac{S(x, R, Q)}{N} + \sum_{\substack{n \leq x \\ \Omega(n; Q) \leq 2N}} \Omega(n; Q) M_{f_A}(n; R)^{2k} \sum_{p^j \parallel n, p > Q} \frac{j \log p}{\log x} \\
&\leq \frac{S(x, R, Q)}{N} + \sum_{\substack{p^j \leq x \\ p > Q}} \frac{j \log p}{\log x} \sum_{\substack{m \leq x/p^j, p \nmid m \\ \Omega(mp^j; Q) \leq 2N}} \Omega(m; Q) M_{f_A}(mp^j; R)^{2k}.
\end{aligned}$$

Combining the above inequality and (10.12), we deduce that

$$T \leq \frac{S(x, R, Q)}{N} + \frac{N+1}{N} \sum_{p^j \leq x} \frac{j^2 \log p}{\log x} \sum_{m \leq x/p^j, p \nmid m} \Omega(m; Q) M_{f_A}(mp^j; R)^{2k} + O_N \left(\frac{x \log Q}{(\log R)^2} \right).$$

We thus conclude that

$$S(x, R, Q) \leq \frac{N+1}{N-1} \sum_{p^j \leq x} \frac{j^2 \log p}{\log x} \sum_{m \leq x/p^j, p \nmid m} \Omega(m; Q) M_{f_A}(mp^j; R)^{2k} + O_N \left(\frac{x \log Q}{(\log R)^2} \right).$$

We can bound the sum on the right hand side in an entirely analogous way to the proof of (10.8). Choosing N sufficiently large, and then C' large enough in terms of N completes the inductive step and thus the proof of (10.11). This proves part (a) of the theorem.

(b) The proof of part (b) is, for the most part, similar to the proof of (10.8). An important detail is that, after applying the induction hypothesis, we use the fact that

$$(10.13) \quad \sum_{p < R} \frac{v \log p}{\log x} \left(\frac{\log(R/p)}{\log R} \right)^{2kA} \cdot \frac{(\log(R/p))^{v \binom{2k}{k} - 2k(A+1)} (\log \log(R/p))^D}{p} \\ \sim \frac{v}{v \binom{2k}{k} - 2k + 1} \cdot \frac{\log R}{\log x} \cdot (\log R)^{\binom{2k}{k} - 2k(A+1)} (\log \log R)^D$$

for any fixed $D \geq 0$, as $R \rightarrow \infty$. This is sufficient when $k \geq 2$, because

$$\frac{v}{v \binom{2k}{k} - 2k + 1} \leq \frac{1 - 1/\binom{2k}{k}}{\binom{2k}{k} - 2k} \leq \frac{1}{2}$$

in this case.

However, when $k = 1$, the situation is more tricky. First of all, we note that the above argument allows to establish the theorem for $x \geq R^{2v/(2v-1)}$. (Note that if $p < R$ and $x \geq R^{2v/(2v-1)}$, then we also have that $x/p \geq (R/p)^{2v/(2v-1)}$, so that the inductive hypothesis can be applied.) Finally, it remains to treat the case when $x \leq R^{2v/(2v-1)}$. We then observe that $\Omega(n; R^\delta, R) \ll_{\epsilon, \delta} 1$, for any fixed δ . It would thus suffice to prove that

$$(10.14) \quad \sum_{n \leq x} v^{\Omega(n; R^\delta)} \delta^{\Omega(n; R^\delta, R)} M_{f_0}(n; R)^2 \leq C'' x (\log R)^{2v-2} (\log \log R)^D \quad (x \geq R \geq 2),$$

for some appropriate constants $C'', D > 0$. Then, in place of (10.13), we note that

$$\sum_{p < R} \frac{(v \cdot \mathbf{1}_{p \leq R^\delta} + \delta \cdot \mathbf{1}_{R^\delta < p \leq R}) \log p}{\log x} \left(\frac{\log(R/p)}{\log R} \right)^{2kA} \cdot \frac{(\log(R/p))^{v \binom{2k}{k} - 2k(A+1)} (\log \log(R/p))^D}{p} \\ < \frac{1}{2} \cdot (\log R)^{\binom{2k}{k} - 2k(A+1)} (\log \log R)^D$$

as $R \rightarrow \infty$, as long as δ is small enough. This allows us to complete the inductive step and establish (10.14), thus completing the proof of the theorem. \square

Given the above result, proving Theorem 1.6 is quite easy:

Proof of Theorem 1.6. (a) This is an immediate consequence of Theorem 10.2(a).

(b) We use Rankin's trick: Given a small $\alpha > 0$ and $1 \leq v \leq 3/2$, we have that

$$\sum_{\substack{n \leq x \\ \Omega(n; R) > \binom{2k}{k} (1+\alpha) \log \log R}} M_{f_A}(n; R)^{2k} \leq \sum_{n \leq x} v^{\Omega(n; R) - \binom{2k}{k} (1+\alpha) \log \log R} M_{f_A}(n; R)^{2k}.$$

Theorem 10.2(b) then implies that

$$\sum_{\substack{n \leq x \\ \Omega(n; R) > \binom{2k}{k}(1+\alpha) \log \log R}} M_f(n; R)^{2k} \ll (\log R)^{\binom{2k}{k} - 2k(A+1) + \binom{2k}{k}(v-1-(1+\alpha) \log v)},$$

provided that v is close enough to 1. We optimize this by choosing $v = 1 + \alpha$, so that $1 - v + (1 + \alpha) \log v = \int_1^{1+\alpha} (\log t) dt > 0$.

Similarly, we have that

$$\sum_{\substack{n \leq x \\ \Omega(n; R) < \binom{2k}{k}(1-\alpha) \log \log R}} M_{f_A}(n; R)^{2k} \leq \sum_{n \leq x} v^{\Omega(n; R) - \binom{2k}{k}(1-\alpha) \log \log R} M_{f_A}(n; R)^{2k},$$

for any $1 - 1/\binom{2k}{k} + \epsilon \cdot \mathbf{1}_{k=1} v \leq 1$. Applying Theorem 10.2(b) and choosing $v = 1 - \alpha$ for small enough α completes the proof of the theorem. \square

10.3. Estimates for general weight functions. It is not so hard to go from estimates for the moments of $M_{f_A}(n; R)$ to the moments of $M_f(n; R)$ for a general weight function f . The following lemma provides the key link.

Lemma 10.3. *Let $f : \mathbb{R} \rightarrow \mathbb{R}$ be supported in $(-\infty, 1]$. Assume further that $f \in C^A(\mathbb{R})$ and that all functions $f, f', \dots, f^{(A)}$ are uniformly bounded for some integer $A \geq 1$. Then*

$$M_f(n; R)^{2k} \ll_{A,k,f} \int_{\frac{\log 2}{\log R}}^1 u^{2k(A-1)} M_{f_{A-1}}(n; R^u)^{2k} du + \frac{1}{(\log R)^{2kA}}.$$

Proof. Since $f(x) = 0$ for $x > 0$ and $f \in C^A(\mathbb{R})$, we must have that $f^{(j)}(1) = 0$ for $j \leq A$. Taylor's theorem with the integral form of the remainder term implies that

$$f(x) = \int_1^x \frac{f^{(A)}(u)}{(A-1)!} (x-u)^{A-1} du = \frac{(-1)^A}{(A-1)!} \int_{[x,1]} f^{(A)}(u) (u-x)^{A-1} du,$$

for all x , since both sides vanish if $x > 1$. Therefore,

$$\begin{aligned} M_f(n; R) &= \frac{(-1)^A}{(A-1)!} \sum_{d|n} \mu(d) \int_{\frac{\log d}{\log R} \leq u \leq 1} f^{(A)}(u) \left(u - \frac{\log d}{\log R}\right)^{A-1} du \\ &= \frac{(-1)^A}{(A-1)!} \int_0^1 f^{(A)}(u) \sum_{\substack{d|n \\ d \leq R^u}} \mu(d) \left(u - \frac{\log d}{\log R}\right)^{A-1} du \\ &= \frac{(-1)^A}{(A-1)!} \int_{\frac{\log 2}{\log R}}^1 f^{(A)}(u) u^{A-1} M_{f_{A-1}}(n; R^u) du + O\left(\frac{1}{(\log R)^A}\right), \end{aligned}$$

by noting that $d = 1$ if $d \leq R^u < 2$. Hölder's inequality then completes the proof. \square

We use the above lemma to show the analogue of Theorem 10.2 for general weight functions f .

Theorem 10.4. *Let $k \in \mathbb{Z}_{\geq 1}$, $x \geq R \geq 2$ and $f : \mathbb{R} \rightarrow \mathbb{R}$ be supported in $(-\infty, 1]$. Assume further that $f \in C^A(\mathbb{R})$ and that all functions $f, f', \dots, f^{(A)}$ are uniformly bounded for some integer $A \geq 1$, and fix some $\epsilon \in (0, 1)$.*

(a) Let $A > \frac{1}{2k} \binom{2k}{k}$. Uniformly for $\eta \in [\log 2 / \log R, 1]$, we have

$$\frac{1}{x} \sum_{n \leq x} \Omega(n; R^\eta) M_f(n; R)^{2k} \ll \frac{\eta}{\log R}.$$

(b) If $A \leq \frac{1}{2k} \binom{2k}{k}$ and $1 - 1/\binom{2k}{k} + \epsilon \cdot \mathbf{1}_{k=1} \leq v \leq 2 - \epsilon$, then

$$\frac{1}{x} \sum_{n \leq x} v^{\Omega(n; R)} M_f(n; R)^{2k} \ll (\log R)^{v \binom{2k}{k} - 2kA} (\log \log R)^{O(1)}.$$

All implied constants depend at most on k , f and ϵ .

Proof. (a) Lemma 10.3 implies that

$$(10.15) \quad \sum_{n \leq x} \Omega(n; R^\eta) M_f(n; R)^{2k} \ll \int_{\frac{\log 2}{\log R}}^1 u^{2k(A-1)} \sum_{n \leq x} \Omega(n; R^\eta) M_{f_{A-1}}(n; R^u)^{2k} du + O\left(\frac{x \log \log R}{(\log R)^{2kA}}\right),$$

When $u \geq \eta$, Theorem 10.2(a) implies that

$$(10.16) \quad \int_{\eta}^1 u^{2k(A-1)} \sum_{n \leq x} \Omega(n; R^\eta) M_{f_{A-1}}(n; R^u)^{2k} du \ll \frac{\eta x}{\log R}.$$

Finally, we consider the integral over $u \in [\log 2 / \log R, \eta]$. Observe that

$$\sum_{n \leq x} \Omega(n; R^u) M_{f_{A-1}}(n; R^u)^{2k} \ll \frac{x}{u \log R},$$

by Theorem 10.2(a). If $x \leq R^{100}$, then we also have that $\Omega(n; R^u, R^\eta) \leq 100/u$ for each $n \leq x$, so that

$$\sum_{n \leq x} \Omega(n; R^u, R^\eta) M_{f_{A-1}}(n; R^u)^{2k} \ll \frac{x}{u^2 \log R}$$

by Theorem 10.2. We thus find that

$$\int_{\frac{\log 2}{\log R}}^{\eta} u^{2k(A-1)} \sum_{n \leq x} \Omega(n; R^\eta) M_{f_{A-1}}(n; R^u)^{2k} du \ll \int_{\frac{\log 2}{\log R}}^{\eta} \frac{u^{2k(A-1)-2} x}{\log R} du \ll \frac{\eta x}{\log R}.$$

Together with (10.15) and (10.16), this proves part (a) in the case when $x \leq R^{100}$.

Finally, let us consider the case when $x > R^{100}$. Then

$$\sum_{n \leq x} \Omega(n; R^u, R^\eta) M_{f_{A-1}}(n; R^u)^{2k} \leq \sum_{\substack{p^j \leq x \\ R^u < p \leq R^\eta}} j \sum_{m \leq x/p^j} M_{f_{A-1}}(m; R^u)^{2k}.$$

When $p^{j-1} \geq \sqrt{x}$, then we bound the inner sum trivially by $\ll (x/p^j)(\log R)^{O(1)}$, so that the total contribution of such summands is $\ll \sqrt{x}(\log R)^{O(1)}$. Finally, when $p^{j-1} \leq \sqrt{x}$, then we see that $x/p^j \geq \sqrt{x}/R \geq x^{0.499} \geq R^u$, so that the sum over m is $\ll (x/p^j)/\log(R^u)$, by Theorem 10.2(a). We thus conclude that

$$\sum_{n \leq x} \Omega(n; R^u, R^\eta) M_{f_{A-1}}(n; R^u)^{2k} \ll \frac{x \log(\eta/u)}{u \log R} + \frac{x}{(u \log R)^2} \ll \frac{x}{u \log R}.$$

Therefore,

$$\int_{\frac{\log 2}{\log R}}^{\eta} u^{2k(A-1)} \sum_{n \leq x} \Omega(n; R^\eta) M_{f_{A-1}}(n; R^u)^{2k} du \ll \frac{\eta x}{\log R}.$$

in this case as well, thus completing the proof of part (a) of the theorem.

(b) This proof of this part is similar. We start again with Lemma 10.3 to find that Lemma 10.3 implies that

$$(10.17) \quad \sum_{n \leq x} v^{\Omega(n; R)} M_f(n; R)^{2k} \ll \int_{\frac{\log 2}{\log R}}^1 u^{2k(A-1)} \sum_{n \leq x} v^{\Omega(n; R)} M_{f_{A-1}}(n; R^u)^{2k} du + O(x(\log R)^{v-1-2kA}),$$

where we also used Theorem III.3.5 in [25, p.308]. Next, if $w = \max\{v, 1\}$ and $\log 2 / \log R \leq u \leq 1$, then note that

$$\sum_{n \leq x} v^{\Omega(n; R)} M_{f_{A-1}}(n; R^u)^{2k} \leq \sum_{\substack{a \leq x \\ P^+(a) \leq R^u}} v^{\Omega(a; R^u)} M_{f_{A-1}}(a; R^u)^{2k} \sum_{\substack{b \leq x/a \\ P^-(b) > R^u}} w^{\Omega(b; R^u, R)}.$$

When $x/a \geq R^u$, the sum over b is $\ll u^{1-w} x / (a \log(R^u))$ by Theorem III.3.5 in [25, p.308]. Hence,

$$\begin{aligned} \sum_{n \leq x} v^{\Omega(n; R)} M_{f_{A-1}}(n; R^u)^{2k} &\ll \frac{u^{1-w} x}{\log(R^u)} \sum_{\substack{a \leq x \\ P^+(a) \leq R^u}} \frac{v^{\Omega(a; R^u)} M_{f_{A-1}}(a; R^u)^{2k}}{a} \\ &\quad + \sum_{\substack{a \leq x \\ P^+(a) \leq R^u}} v^{\Omega(a; R^u)} M_{f_{A-1}}(a; R^u)^{2k}. \end{aligned}$$

We bound the first sum by Theorem 10.1(b) and the second one by Theorem 10.2(b) to find that

$$\sum_{n \leq x} v^{\Omega(n; R)} M_{f_{A-1}}(n; R^u)^{2k} \ll xu^{1-w+v\binom{2k}{k}-2kA} (\log R)^{v\binom{2k}{k}-2kA}.$$

Inserting the above bound into (10.17) completes the proof of the theorem. \square

We conclude this section with the proof of Theorem 1.4. We need a preliminary lemma.

Lemma 10.5. *If $m \in \mathbb{Z}_{\geq 1}$, $g \in C^1(\mathbb{R}^m)$ and $z \geq y \geq 3$, then there is a positive constant $c > 0$ such that*

$$\begin{aligned} \sum_{y < p_1 < \dots < p_m \leq z} \frac{g(\log p_1, \dots, \log p_m)}{(p_1 - 1) \cdots (p_m - 1)} &= \frac{1}{m!} \int_{[\log y, \log z]^m} \frac{g(t_1, \dots, t_m)}{t_1 \cdots t_m} dt_1 \cdots dt_m \\ &\quad + O\left(\frac{\|g\|_\infty + \|\nabla g\|_\infty}{e^{c\sqrt{\log y}}} \cdot \frac{(\sum_{y < p \leq z} 1/p)^{m-1} + (\int_y^z dt/t \log t)^{m-1}}{m!/m^2} \right). \end{aligned}$$

Proof. First of all, note that

$$\begin{aligned} \sum_{y < p_1 < \dots < p_m \leq z} \frac{g(\log p_1, \dots, \log p_m)}{(p_1 - 1) \cdots (p_m - 1)} &= \frac{1}{m!} \sum_{\substack{y < p_1, \dots, p_m \leq z \\ \text{distinct}}} \frac{g(\log p_1, \dots, \log p_m)}{(p_1 - 1) \cdots (p_m - 1)} \\ &= \frac{1}{m!} \sum_{y < p_1, \dots, p_m \leq z} \frac{g(\log p_1, \dots, \log p_m)}{p_1 \cdots p_m} \\ &\quad + O\left(\frac{m^2 \|g\|_\infty (\sum_{y < p \leq z} 1/p)^{m-1}}{m! y}\right). \end{aligned}$$

So, if we can show that

$$\begin{aligned} \sum_{y < p_1, \dots, p_m \leq z} \frac{g(\log p_1, \dots, \log p_m)}{p_1 \cdots p_m} &= \int_{[\log y, \log z]^m} \frac{g(t_1, \dots, t_m)}{t_1 \cdots t_m} dt_1 \cdots dt_m \\ &\quad + O\left(\sum_{j=1}^m \frac{\|g\|_\infty + \|\partial g / \partial x_j\|_\infty}{e^{c\sqrt{\log y}}} \left(\int_y^z \frac{dt}{t \log t}\right)^{j-1} \left(\sum_{y < p \leq z} \frac{1}{p}\right)^{m-j}\right), \end{aligned}$$

then the lemma will follow. But this estimate can be easily proved by induction on m and the Prime Number Theorem, and the proof is completed. \square

Let us now see how we can deduce Theorem 1.4 from the above results:

Proof of Theorem 1.4. The first estimate of part (a) follows by 10.4(a) and part (b) follows by Theorem 10.2(b). It remains to prove the second estimate of part (a). Note that it suffices to prove that

$$\prod_{p \leq R} \left(1 - \frac{1}{p}\right) \sum_{P^+(n) \leq R} \frac{M_f(n; R)^{2k}}{n} = \frac{c_{k,f}}{\log R} + O\left(\frac{1}{(\log R)^{2-\epsilon}}\right),$$

since $x \geq R^{2k} \log^2 R$ here. Fix $\eta \in [\log 2 / \log R, 1]$ to be chosen later. Then Theorem 10.4(a) implies that

$$\sum_{\substack{P^+(n) \leq R \\ P^-(n) \leq R^\eta}} \frac{M_f(n; R)^{2k}}{n} \leq \sum_{P^+(n) \leq R} \frac{\Omega(n; R^\eta) M_f(n; R)^{2k}}{n} \ll \eta.$$

Moreover, for each positive integer m , we have that

$$\sum_{\substack{p|n \Rightarrow R^\eta < p \leq R \\ \omega(n)=m}} \frac{M_f(n; R)^{2k}}{n} = \sum_{R^\eta < p_1 < \dots < p_m \leq R} \frac{M_f(p_1 \cdots p_m; R)}{(p_1 - 1) \cdots (p_m - 1)}$$

We note that

$$M_f(p_1 \cdots p_m; R) = g_m\left(\frac{\log p_1}{\log R}, \dots, \frac{\log p_m}{\log R}\right),$$

where

$$g_m(t_1, \dots, t_m) := \sum_{J \subset \{1, \dots, m\}} (-1)^{\#J} f\left(\sum_{j \in J} t_j\right)$$

which is a smooth function satisfying the estimates $\|g_m^{2k}\|_\infty \leq 4^{km}\|f\|_\infty^{2k}$ and $\|\nabla g_m^{2k}\|_\infty \leq 2k4^{km}\|f'\|_\infty\|f\|_\infty^{2k-1}$. Therefore Lemma 10.5 implies that

$$\sum_{\substack{p|n \Rightarrow R^\eta < p \leq R \\ \omega(n)=m}} \frac{M_f(n; R)^{2k}}{n} = \frac{1}{m!} \int_{[\eta, 1]^m} \frac{g_m(t_1, \dots, t_m)^{2k}}{t_1 \cdots t_m} dt_1 \cdots dt_m + O\left(\frac{(4^k \log(1/\eta) + O(1))^m}{e^{c\sqrt{\log(R^\eta)}} m! / m^2}\right)$$

for all $m \in \mathbb{Z}_{\geq 1}$. We thus conclude that

$$(10.18) \quad \sum_{P^+(n) \leq R} \frac{M(n; R)^{2k}}{n} = F(\eta) + O\left(\eta + \frac{\log^2(1/\eta)}{\eta^{4^k} e^{c\sqrt{\log(R^\eta)}}}\right) \quad (0 < \eta \leq 1, R^\eta \geq 2),$$

where

$$F(\eta) := 1 + \sum_{m=1}^{\infty} \frac{1}{m!} \int_{[\eta, 1]^m} \frac{g_m(t_1, \dots, t_m)^{2k}}{t_1 \cdots t_m} dt_1 \cdots dt_m.$$

Completing the proof is now an exercise in real analysis. We start by proving that $\lim_{\eta \rightarrow 0^+} F(\eta)$ exists. Indeed, applying (10.18) twice, we deduce that

$$F(\eta_1) - F(\eta_2) \ll \max_{j \in \{1, 2\}} \left\{ \eta_j + \frac{\log^2(1/\eta_j)}{\eta_j^{4^k} e^{c\sqrt{\log(R^{\eta_j})}}}\right\}$$

whenever $0 < \eta_1 \leq \eta_2 \leq 1$ and $R^{\eta_1} \geq 2$. Letting $R \rightarrow \infty$, we find that

$$(10.19) \quad F(\eta_1) - F(\eta_2) \ll \eta_2 \quad (0 < \eta_1 \leq \eta_2 \leq 1).$$

In particular, Cauchy's convergence criterion implies that $\lim_{\eta \rightarrow 0^+} F(\eta)$ exists. Call F this limit, which clearly equals $F(0)$, and note that letting $\eta_1 \rightarrow 0^+$ in (10.19) implies that

$$F(\eta) = F + O(\eta) \quad (0 < \eta \leq 1).$$

Together with (10.18), this yields the estimate

$$\sum_{P^+(n) \leq R} \frac{M(n; R)^{2k}}{n} = F + O\left(\eta + \frac{\log^2(1/\eta)}{\eta^{4^k} e^{c\sqrt{\log(R^\eta)}}}\right) \quad (0 < \eta \leq 1, R^\eta \geq 2).$$

Selecting η such that

$$R^\eta = e^{(\log \log R)^3}$$

completes the proof of Theorem 1.4. □

We conclude this section with the proof of Theorem 1.7.

Proof of Theorem 1.7. The first estimate of part (a) can be proven following *mutatis mutandis* the proof of Theorem 1.6(a) above, using Theorem 10.4 in place of Theorem 10.2. Similarly, part (b) follows from the proof of Theorem 1.6(b). □

11. THE ANALOGY FOR NON-EXCEPTIONAL DIRICHLET CHARACTERS

In the section, we study the sum

$$\mathcal{X}_{2k}(R) = \prod_{p \leq R} \left(1 - \frac{1}{p}\right) \sum_{P^+(n) \leq R} \frac{1}{n} \left(\sum_{\substack{d|n \\ R/2 < d \leq R}} \chi(d) \right)^{2k}$$

when $k \geq 2$ and $L(1, \chi)$ is not very small and prove Theorem 1.5(b).

We have that

$$\mathcal{X}_{2k}(R) = \sum_{R/2 < d_1, \dots, d_{2k} \leq R} \frac{\chi(d_1)\chi(d_2) \cdots \chi(d_{2k})}{[d_1, \dots, d_{2k}]}.$$

We want to introduce new variables D_I , $I \in \mathcal{S}^*(2k)$, as in Section 2, but first we perform a technical manoeuvre to simplify the situation. We write $d_i = d'_i d''_i$, where $P^+(d'_i) \leq y < P^-(d''_i)$, where

$$y := (\log R)^{4k+1}.$$

The contribution to $\mathcal{X}_{2k}(R)$ of \mathbf{d} 's for which d''_i is not square-free for some i is $\ll (\log R)^{4k-1}/y$ by a crude upper bound, and so is the contribution of those \mathbf{d} 's with $\max_i d'_i > B$, where

$$B := e^{(\log \log R)^3},$$

by Rankin's trick. Then we let D_I , $I \in \mathcal{S}^*(2k)$, be the product of those primes that divide d''_i when $i \in I$, and are coprime to the other d''_i 's. The numbers D_I are pairwise coprime and square-free, and $d''_i = \prod_{I \in \mathcal{S}^*(2k), I \ni i} D_I$, so that

$$\frac{\chi(d''_1)\chi(d''_2) \cdots \chi(d''_{2k})}{[d''_1, \dots, d''_{2k}]} = \frac{\prod_{I \in \mathcal{S}^+(2k)} \chi_0(D_I) \prod_{I \in \mathcal{S}^-(2k)} \chi(D_I)}{\prod_{I \in \mathcal{S}^*(2k)} D_I}.$$

We may now drop the condition that the D_I 's are square-free and coprime, since the contribution to $\mathcal{I}_{2k}(R)$ of the D_I 's not satisfying these conditions is $\ll (\log R)^{4k-1}/y$. Finally, we may drop the condition that $(D_I, q) = 1$ for $I \in \mathcal{S}^+(2k)$, encoded in the notation $\chi_0(D_I)$, since the contribution of D_I 's with $P^-(D_I) > y$ and $(D_I, q) > 1$ is $\ll (\log R)^{4k-1} \sum_{p|q, p > y} 1/p \ll (\log R)^{4k}/y$. The above discussion implies that

$$\begin{aligned} \mathcal{X}_{2k}(R) &= \sum_{\substack{P^+(d'_i) \leq y, d'_i \leq B \\ 1 \leq i \leq 2k}} \frac{\chi(d'_1) \cdots \chi(d'_{2k})}{[d'_1, \dots, d'_{2k}]} \sum_{\substack{P^-(D_I) > y \ (I \in \mathcal{S}^*(2k)) \\ R/(2d'_i) < \prod_{I \in \mathcal{S}^*(2k), I \ni i} D_I \leq R/d'_i \\ 1 \leq i \leq 2k}} \frac{\prod_{I \in \mathcal{S}^-(2k)} \chi(D_I)}{\prod_{I \in \mathcal{S}^*(2k)} D_I} \\ &\quad + O\left(\frac{1}{\log R}\right). \end{aligned}$$

Next, we note that

$$\sum_{\substack{n \leq x \\ P^-(n) > y}} \chi(n) \ll x^{1-1/(30 \log y)} \quad (x \geq \max\{q^4, y\}),$$

by Lemma 2.4 in [16]. Therefore,

$$(11.1) \quad \sum_{\substack{n > q^4 B \\ P^-(n) > y}} \frac{\chi(n)}{n} \ll \frac{1}{y}.$$

This implies that the contribution to $\mathcal{X}_{2k}(R)$ with $D_I > q^4 B$ for some $I \in \mathcal{S}^-(2k)$ is $\ll (\log R)^{4k-1}/y$. To conclude, we have shown

$$(11.2) \quad \begin{aligned} \mathcal{X}_{2k}(R) = & \sum_{\substack{P^+(d'_i) \leq y, d'_i \leq B \\ 1 \leq i \leq 2k}} \sum_{\substack{D_I \leq q^4 B \\ P^-(D_I) > y \\ I \in \mathcal{S}^-(2k)}} \frac{\chi(d'_1) \cdots \chi(d'_{2k})}{[d'_1, \dots, d'_{2k}]} \cdot \frac{\prod_{I \in \mathcal{S}^-(2k)} \chi(D_I)}{\prod_{I \in \mathcal{S}^-(2k)} D_I} \cdot T(R_1, \dots, R_{2k}) \\ & + O\left(\frac{1}{\log R}\right). \end{aligned}$$

where $R_i = R/(d'_i \prod_{I \in \mathcal{S}^-(2k), I \ni i} D_I)$ and

$$T(\mathbf{R}) := \sum_{\substack{P^-(D_I) > y \ (I \in \mathcal{S}^+(2k)) \\ R_i/2 < \prod_{I \in \mathcal{S}^+(2k), I \ni i} D_I \leq R_i \\ 1 \leq i \leq 2k}} \frac{1}{\prod_{I \in \mathcal{S}^+(2k)} D_I}.$$

Our task now becomes estimating $T(\mathbf{R})$. Let $d = 2^{2k-1} - 2k$ and recall the definition of $V_d(\cdot)$ from the statement of Theorem 1.5(b). The proof of Theorem 1.1 shows that $V_k(m) \asymp m^d$. We claim that

$$(11.3) \quad T(\mathbf{R}) = V_k(\log R) \left(1 + O\left(\frac{\log B}{\log R}\right)\right) \prod_{p \leq y} \left(1 - \frac{1}{p}\right)^{2^{2k-1}-1}$$

whenever $R/B \leq R_i \leq R$ for all i , as is the case here. Proving (11.3) can be accomplished easily using a lattice point count and the fundamental lemma of sieve methods. First of all, note that the part of $T(\mathbf{R})$ where $D_I \leq B$ for some $I \in \mathcal{S}^+(2k)$ is trivially $\ll (\varphi(P)/P)^{2^{2k-1}-1} (\log R)^{d-1} \log B$ by an upper bound sieve, where we have set $P := \prod_{p \leq y} p$ for simplicity. In the rest of the range, we set $\rho = 1 + 1/\log R$ and divide the variables D_I into boxes of the form $(\rho^{m_I}, \rho^{m_I+1}]$, $m_I \geq 0$. Replacing D_I by ρ^{m_I} in the conditions $R_i/2 < \prod_{J \in \mathcal{S}^+(2k), J \ni i} D_J \leq R_i$ creates a total error of size $\ll (\varphi(P)/P)^{2^{2k-1}-1} (\log R)^{d-1}$. In addition, if $\rho^{m_I} \geq B = e^{(\log \log R)^3}$, then we have that

$$\sum_{\substack{\rho^{m_I} < D_I \leq \rho^{m_I+1} \\ P^-(D_I) > y}} \frac{1}{D_I} = (\log \rho) \frac{\varphi(P)}{P} + O_C\left(\frac{1}{(\log R)^C}\right)$$

for any fixed C , by the fundamental lemma of sieve methods (see, for example, [25, Theorem I.4.3]). We thus conclude that

$$\begin{aligned} T(\mathbf{R}) = & ((\log \rho) \varphi(P)/P)^{2^{2k-1}-1} \sum_{\substack{m_I \geq \log B / \log \rho \ (I \in \mathcal{S}^+(2k)) \\ \frac{\log(R_i/2)}{\log \rho} < \sum_{I \in \mathcal{S}^+(2k), I \ni i} m_I \leq \frac{\log R_i}{\log \rho}}} 1 \\ & + O\left((\varphi(P)/P)^{2^{2k-1}-1} (\log R)^{d-1} \log B\right). \end{aligned}$$

A straightforward lattice point counting argument implies that the sum on the right hand side equals

$$\frac{W_k(\log R_1, \dots, \log R_{2k})}{(\log \rho)^{2^{2k-1}-1}} \left(1 + O\left(\frac{\log B}{\log R}\right)\right),$$

where $W_k(\mathbf{m})$ is the volume of the polytope $\{(x_I)_{I \in S^+(2k)} : x_I \geq 0 \forall I, m_i - \log 2 \leq \sum_{I \ni i} x_I \leq m_i \forall i\}$. Since $m_i = \log R_i = \log R + O(\log B)$ here, we may show using the Mean Value Theorem that $W_k(\mathbf{m}) = V_k(\mathbf{m}) + O((\log R)^{d-1} \log B)$. Relation (11.3) then follows.

We are now ready to complete the proof of Theorem 1.5(a): inserting the estimate (11.3) into (11.2), we conclude that

$$\begin{aligned} \mathcal{X}_{2k}(R) &= \frac{V_k(\log R)}{(P/\varphi(P))^{2^{2k-1}-1}} \sum_{\substack{P^+(d'_i) \leq y, d_i \leq B \\ 1 \leq i \leq 2k}} \sum_{\substack{D_I \leq q^4 B \\ P^-(D_I) > y \\ I \in S^-(2k)}} \frac{\chi(d'_1) \cdots \chi(d'_{2k})}{[d'_1, \dots, d'_{2k}]} \cdot \frac{\prod_{I \in S^-(2k)} \chi(D_I)}{\prod_{I \in S^-(2k)} D_I} \\ &\quad + O((\log R)^{d-1} (\log(q \log R))^{O(1)}). \end{aligned}$$

We now remove the conditions $D_I \leq q^4 B$ and $d'_i \leq B$ via (11.1) and Rankin's trick, respectively. We conclude that

$$\begin{aligned} \mathcal{X}_{2k}(R) &= \frac{V_k(\log R)}{(P/\varphi(P))^{2^{2k-1}-1}} \sum_{\substack{P^+(d'_i) \leq y \\ 1 \leq i \leq 2k}} \sum_{\substack{P^-(D_I) > y \\ I \in S^-(2k)}} \frac{\chi(d'_1) \cdots \chi(d'_{2k})}{[d'_1, \dots, d'_{2k}]} \cdot \frac{\prod_{I \in S^-(2k)} \chi(D_I)}{\prod_{I \in S^-(2k)} D_I} \\ &\quad + O((\log R)^{d-1} (\log(q \log R))^{O(1)}). \end{aligned}$$

Finally, we note that

$$\sum_{\substack{P^+(d'_i) \leq y \\ 1 \leq i \leq 2k}} \frac{\chi(d'_1) \cdots \chi(d'_{2k})}{[d'_1, \dots, d'_{2k}]} = \prod_{p \leq y} \left(1 + \sum_{j=1}^{\infty} \sum_{\substack{j_1, \dots, j_{2k} \geq 0 \\ \max\{j_1, \dots, j_{2k}\} = j}} \frac{\chi(p)^{j_1 + \dots + j_{2k}}}{p^j} \right).$$

The coefficient of $1/p$ is $(2^{2k-1} - 1)\chi_0(p) + 2^{2k-1}\chi(p)$. We thus conclude that

$$\begin{aligned} &(\varphi(P)/P)^{2^{2k-1}-1} \sum_{\substack{P^+(d'_i) \leq y \\ 1 \leq i \leq 2k}} \sum_{\substack{P^-(D_I) > y \\ I \in S^-(2k)}} \frac{\chi(d'_1) \cdots \chi(d'_{2k})}{[d'_1, \dots, d'_{2k}]} \cdot \frac{\prod_{I \in S^-(2k)} \chi(D_I)}{\prod_{I \in S^-(2k)} D_I} \\ &= \prod_{p \leq y} \left(1 + \sum_{j=1}^{\infty} \sum_{\substack{j_1, \dots, j_{2k} \geq 0 \\ \max\{j_1, \dots, j_{2k}\} = j}} \frac{\chi(p)^{j_1 + \dots + j_{2k}}}{p^j} \right) \left(1 - \frac{1}{p} \right)^{2^{2k-1}-1} \prod_{p > y} \left(1 - \frac{\chi(p)}{p} \right)^{-2^{2k-1}} \\ &= \prod_p \left(1 + \sum_{j=1}^{\infty} \sum_{\substack{j_1, \dots, j_{2k} \geq 0 \\ \max\{j_1, \dots, j_{2k}\} = j}} \frac{\chi(p)^{j_1 + \dots + j_{2k}}}{p^j} \right) \left(1 - \frac{1}{p} \right)^{2^{2k-1}-1} + O\left(\frac{1}{y}\right). \end{aligned}$$

An easy calculation then completes the proof of Theorem 1.5(b).

12. THE ANALOGY FOR EXCEPTIONAL DIRICHLET CHARACTERS

In this section, we consider the quantity $\mathcal{X}_{2k}(R)$ when $k = 1$, or when when $\chi(p) = -1$ for most primes $p \leq R$, and complete the proof of Theorem 1.5. Our arguments here resemble closely the ones of Section 8, so we only highlight the key points here. Throughout the proof, we assume that $R \geq q^{2c_1}$, the complimentary case being trivial.

12.1. Initial preparations. Arguing as in Section 8.4, we find that

$$\begin{aligned} \mathcal{X}_{2k}(R) &= \frac{1}{(2i\pi)^{2k}} \int \cdots \int \sum_{\substack{\operatorname{Re}(s_j) = \lambda^j / \log R \\ |\operatorname{Im}(s_j)| \leq T \\ 1 \leq j \leq 2k}} \frac{\prod_{j=1}^{2k} \chi(d_j) d_j^{-s_j}}{[d_1, \dots, d_{2k}]} \prod_{j=1}^{2k} \widehat{h}_R(s_j) (1 - 2^{-s_j}) ds_1 \cdots ds_{2k} \\ &\quad + O\left(\frac{1}{(\log R)^{100}}\right), \end{aligned}$$

where h is a smooth function with $h(x) = 1$ for $x \leq 1 - 1/(\log R)^{(2k-1)2^{2k+1}+200k+2}$ and $h(x) = 0$ for $x > 1$, $T = \exp\{(\log \log R)^2\}$, and λ is some large parameter > 1 to be chosen later.

By expanding as an Euler product, we find that for $\operatorname{Re}(s_1), \dots, \operatorname{Re}(s_{2k}) \geq -1/4k$ we have

$$\begin{aligned} \sum_{d_1, \dots, d_{2k} \geq 1} \frac{\prod_{j=1}^{2k} \chi(d_j) d_j^{-s_j}}{[d_1, \dots, d_{2k}]} &= \prod_p \left(1 + \sum_{I \in S^*(2k)} \frac{\chi(p)^{\#I}}{p^{1+s_I}} + O\left(\frac{1}{p^{2-2k/4^k}}\right) \right) \\ &= P(\mathbf{s}) \prod_{I \in S^*(2k)} L(1 + s_I, \chi^{\#I}), \end{aligned}$$

where $P(\mathbf{s})$ is given by an Euler product which converges absolutely in the region $\operatorname{Re}(s_j) > -1/4k$ for all j . Next, we set

$$F(\mathbf{s}) = P(\mathbf{s}) \prod_{j=1}^{2k} \frac{\widehat{h}_R(s_j) (1 - 2^{-s_j})}{R^{s_j}}$$

and

$$\zeta_q(s) := L(s, \chi_0) = \zeta(s) \prod_{p|q} \left(1 - \frac{1}{p^s} \right),$$

so that

$$\begin{aligned} \mathcal{X}_{2k}(R) &= \frac{1}{(2i\pi)^{2k}} \int \cdots \int \sum_{\substack{\operatorname{Re}(s_j) = \lambda^j / \log R \\ |\operatorname{Im}(s_j)| \leq T \\ 1 \leq j \leq 2k}} F(\mathbf{s}) R^{s_{[2k]}} \prod_{I \in S^-(2k)} L(1 + s_I, \chi) \\ (12.1) \quad &\quad \times \prod_{I \in S^+(2k)} \zeta_q(1 + s_I) ds_1 \cdots ds_{2k} + O\left(\frac{1}{(\log R)^{100}}\right). \end{aligned}$$

Similarly to Section 8, we let $\widetilde{\mathcal{C}}_\ell$ denote the class of complex-valued functions f defined in a domain containing

$$\widetilde{\Omega}_\ell := \{\mathbf{s} \in \mathbb{C}^\ell : |\operatorname{Re}(s_j)| < 1/5k, |\operatorname{Im}(s_j)| < T + 1 \ (1 \leq j \leq \ell)\},$$

it is analytic in $\tilde{\Omega}_\ell$, and its derivatives satisfy the bound

$$(12.2) \quad \frac{\partial^{j_1+\dots+j_\ell} f}{\partial s_1^{j_1} \dots \partial s_\ell^{j_\ell}}(\mathbf{s}) \ll_{j_1, \dots, j_\ell} \prod_{m=1}^{\ell} \frac{[(1 + (qT)^{-\operatorname{Re}(s_m)})(\log \log R)^{j_m}]^{O(1)}}{|s_m| + 1}$$

for all $j_1, \dots, j_\ell \geq 0$ and all $\mathbf{s} = (s_1, \dots, s_\ell) \in \tilde{\Omega}_\ell$.

Since there is an absolute constant $c_1 > 0$ such that

$$(12.3) \quad L^{(j)}(s, \psi) \ll_m (1 + (q + |t|)^{c_1(1-\sigma)}) \log^{j+1}(q + |t|) + \frac{\mathbf{1}_{\psi=\chi_0}}{|s-1|^{j+1}}$$

for $j \in \mathbb{Z}_{\geq 0}$, $\psi \in \{\chi, \chi_0\}$ and $j \in \{0, 1\}$, a standard consequence of bounds on the exponential sum $\sum_{n \leq N, n \equiv a \pmod{q}} n^{it}$ (see, for example, Lemma 4.1 in [16]), we have that F is in the class $\tilde{\mathcal{C}}_{2k}$.

12.2. The case $k = 1$. We first deal with the case $k = 1$ that is easy and will help us clarify some of the technical details of the argument. When $k = 1$, we move the variable s_2 to the line $\operatorname{Re}(s_2) = -\epsilon$, for a sufficiently small ϵ . The contribution of the horizontal contours is $\ll (\log R)^{O(1)}/T$, and the contribution of the contour $\operatorname{Re}(s_2) = -\delta$ is $\ll R^{-\delta}$, for some positive $\delta = \delta(\epsilon)$ by (12.3), and by our assumptions that $F \in \mathcal{C}_2$ and that $R \geq q^{2c_1}$. In conclusion,

$$\mathcal{X}_2(R) = \frac{1}{2i\pi} \int_{\substack{\operatorname{Re}(s_1)=\lambda/\log R \\ |\operatorname{Im}(s_1)| \leq T \\ 1 \leq j \leq 2k}} F(s_1, -s_1) L(1 + s_1, \chi) L(1 - s_1, \chi) ds_1 + O\left(\frac{1}{(\log R)^{100}}\right).$$

Finally, we move s_1 to the line $\operatorname{Re}(s_1) = 0$. No poles are encountered, and the contribution of the horizontal lines is easily seen to be $\ll (\log R)^{O(1)}/T$, so that

$$\begin{aligned} \mathcal{X}_2(R) &= \frac{1}{2\pi} \int_{-T}^T F(it, -it) |L(1 + it, \chi)|^2 dt + O\left(\frac{1}{(\log R)^{100}}\right) \\ &= \frac{1}{2\pi} \int_{-(\log R)^{102}}^{(\log R)^{102}} P(it, -it) \left| L(1 + it, \chi) \hat{h}_R(it) (1 - 2^{it}) \right|^2 dt + O\left(\frac{1}{(\log R)^{100}}\right), \end{aligned}$$

since $\hat{h}_R(it)(1 - 2^{it}) \ll 1/(1 + |t|)$. Finally, using 6.6 to replace $\hat{h}_R(s)$ by R^s/s , choosing C to be large enough, and then extending the range of integration to \mathbb{R} yields the estimate

$$\mathcal{X}_2(R) = \frac{1}{2\pi} \int_{-\infty}^{\infty} P(it, -it) \left| L(1 + it, \chi) \cdot \frac{\sin(t(\log 2)/2)}{t} \right|^2 dt + O\left(\frac{1}{(\log R)^{100}}\right),$$

which proves Theorem 1.5(a). (Obviously, we can obtain a much stronger error term, but we have chosen to content ourselves with a more qualitative result.)

12.3. Contour shifting. Next, we focus on the case $k \geq 2$ and prove Theorem 1.5(c). From now on, we will always be working under the assumptions and notations

$$L(\beta, \chi) = 0, \quad \beta > 1 - 1/(100 \log q), \quad Q = e^{1/(1-\beta)}.$$

As it is well-known, we have that $\sum_{q < p \leq Q} (1 + \chi(p))/p \ll 1$ and $\sum_{p > Q} \chi(p)/p \ll 1$ (see, for example, Theorems 2.1 and 2.4 in [17]). As a consequence, we note once and for all that

$$(12.4) \quad L(1, \chi) \asymp \frac{1}{\log Q} \prod_{p \leq q} \left(1 + \frac{1 + \chi(p)}{p}\right).$$

As in Section 8, we shift the contours of the variables s_1, \dots, s_{2k} in a certain order. As in that section, to describe the general contour shifting argument after N steps, $0 \leq N \leq 2k$, we fix sets I_1, \dots, I_N , and distinct integers $j_n \in I_n$ for each n . Recall, also, that s_j denotes a variable and x_j denotes a linear form. We then define

$$V_n = \text{Span}_{\mathbb{Q}}(x_{I_1}, \dots, x_{I_n}) \quad \text{and} \quad \mathcal{I}_N = \{I \in \mathcal{S}(2k) : x_I \in V_n\} \quad (0 \leq n \leq N).$$

Imposing the conditions $x_{I_1} = \dots = x_{I_N} = 0$, we may write x_{j_1}, \dots, x_{j_N} in terms of the variables s_j with $j \in [2k] \setminus \{j_1, \dots, j_N\}$. Hence x_I becomes a linear form $L_{N,I}$ in the variables s_j with $j \in [2k] \setminus \{j_1, \dots, j_N\}$. Moreover, $x_I \in V_N$ if and only if $L_{N,I} = 0$.

As we will see, we will always be able to assume that $j_n = 2k - n + 1$. Let $d \in \mathbb{Z}_{\geq 0}$ and Given the above set-up with $j_n = 2k - n + 1$, an integer $d \in \mathbb{Z}_{\geq 0}$ and $\mathbf{h} = (h_{n,I})_{0 \leq n \leq N, I \in \mathcal{S}^*(2k)}$ be a vector of non-negative integers such that:

- $0 = h_{0,I} \leq h_{1,I} \leq \dots \leq h_{N,I}$ for $I \in \mathcal{S}^*(2k)$;
- If $I \in \mathcal{I}_n \setminus \mathcal{I}_{n-1}$ for some $n \in \{1, \dots, N\}$, then $h_{m,I} = h_{n,I}$ for all $m \geq n$;
- $X_N \geq N + d$, where

$$X_N := \#(\mathcal{I}_N \cap \mathcal{S}^+(2k)) - \sum_{I \in \mathcal{S}^-(2k) \cup (\mathcal{S}^+(2k) \setminus \mathcal{I}_N)} h_{N,I}.$$

A function $J : \mathbb{R}_{\geq 2} \rightarrow \mathbb{C}$ is called a **fundamental component of level N and of type (I, \mathbf{h}, d)** if:

- when $N = 2k$, it equals

$$J(R) = (\log R)^{X_N - N - d} \prod_{I \in \mathcal{S}^-(2k)} L^{(h_{N,I})}(1, \chi);$$

- when $N < 2k$, it is of the form

$$\begin{aligned} J(R) &= \left(\frac{\varphi(q)}{q} \right)^{M_N} \frac{(\log R)^{X_N - N - d}}{(2i\pi)^{2k - N}} \prod_{I \in \mathcal{S}^-(2k) \cap \mathcal{I}_N} L^{(h_{N,I})}(1, \chi) \\ &\quad \times \int \dots \int_{\substack{\text{Re}(s_j) = \lambda_j / \log R \\ |\text{Im}(s_j)| \leq T \\ 1 \leq j \leq 2k - N}} G(\mathbf{s}) R^{E_N(\mathbf{s})} \prod_{I \in \mathcal{S}^-(2k) \setminus \mathcal{I}_N} L^{(h_{N,I})}(1 + L_{N,I}(\mathbf{s}), \chi) \\ &\quad \times \prod_{I \in \mathcal{S}^+(2k) \setminus \mathcal{I}_N} \zeta_q^{(h_{N,I})}(1 + L_{N,I}(\mathbf{s})) ds_{2k-N} \dots ds_1, \end{aligned}$$

where $\lambda_j / \lambda_{j-1} \geq \lambda$, $E_N(\mathbf{s}) := L_{N,[2k]}(\mathbf{s})$,

$$M_N := \sum_{n=1}^N \sum_{I \in (\mathcal{I}_n \setminus \mathcal{I}_{n-1}) \cap \mathcal{S}^+(2k)} (h_{n-1,I} + 1),$$

and G is a function in the variables s_1, \dots, s_{2k-N} that belongs to the class \mathcal{C}_{2k-N} , given by

$$G(\mathbf{s}) = F(L_{N,\{1\}}(\mathbf{s}), \dots, L_{N,\{2k\}}(\mathbf{s}))$$

when $d = 0$. In particular, G is non-vanishing in Ω_{2k-N} when $d = 0$ by (12.3).

As in Section 8.2, a fundamental component of level N is called reducible when $N < 2k$ and $E_N \neq 0$. Otherwise, it is called irreducible. With this above terminology, the integral on the right hand side of (12.1) is a reducible fundamental component of level 0 and of type $(\emptyset, \emptyset, 0)$.

Again as in Section 8.2, when $E_N \neq 0$ there are some $\gamma_j \in \mathbb{Q}$ with $\gamma_{j_{N+1}} \neq 0$ such that

$$E_N(\mathbf{x}) = \gamma_1 x_1 + \gamma_2 x_2 + \cdots + \gamma_{j_{N+1}} x_{j_{N+1}}$$

If λ is big enough, then the sign of $\operatorname{Re}(E_N(\mathbf{s}))$ throughout the region of integration is constant and equal to the sign as $\gamma_{j_{N+1}}$.

The analogies of Lemmas 8.4 and 8.5 can be proven in this setting:

Lemma 12.1. *Assume the above setup, let $J(R)$ be a reducible fundamental component of level $N < 2k$, and let $\gamma_{j_{N+1}}$ be as above. Suppose, further, that $k \geq 2$ and that $e^{(\log q)^2} \leq R \leq Q$. All implied constants below depend at most on k .*

- (a) *If $\gamma_{j_{N+1}} > 0$, then $J(R)$ is a linear combination of fundamental components of level $N+1$, up to an error term of size $\ll 1/\log R$. Moreover, the coefficients of this linear combination are $\ll (\log q)^{O(1)}$.*
- (b) *If $\gamma_{j_{N+1}} < 0$, then $J(R) \ll T^{-1+o(1)}$.*

We iterate the above lemma till all the fundamental components we are dealing with are irreducible. For such components, we have the following asymptotic formula.

Lemma 12.2. *Assume the above setup. Suppose, further, that $k \geq 2$ and that $e^{(\log q)^2} \leq R \leq Q$. If $J(R)$ is an irreducible fundamental component, then there is some complex number $c \ll (\log q)^{O(1)}$ such that*

$$J(R) = c(\log R)^{\binom{2k}{k}-2k} + O((\log(q \log R))^{O(1)}(\log R)^{\binom{2k}{k}-2k-1}).$$

All implied constants depend at most on k .

Since the integral on the right hand side of (12.1) is a reducible fundamental component of level 0, we apply Lemma 8.4 repeatedly to write it as a linear combination of irreducible fundamental components, and then estimate these components by Lemma 8.5. This proves that there is a constant $c_k(\chi) \ll (\log q)^{O(1)}$ such that

$$(12.5) \quad \mathcal{X}_{2k}(R) = c_k(\chi) \cdot (\log R)^{\binom{2k}{k}-2k} + O\left((\log(q \log R))^{O(1)}(\log R)^{\binom{2k}{k}-2k-1}\right)$$

when $k \geq 2$ and $e^{(\log q)^2} \leq R \leq Q$. We will show that $c_k(\chi) \gg (\log q)^{-O(1)}$ in Section 12.5 and complete the proof of Theorem 1.5. The key intermediate Lemmas 12.1 and 12.2 are proven in the next section.

12.4. Proof of the auxiliary Lemmas 12.1 and 12.2. For easy reference, we record the following bound that we will repeatedly use: for $R \leq Q$, we have

$$(12.6) \quad (\log R)^{X_N - N - d} \prod_{I \in \mathcal{S}^-(2k) \cap \mathcal{I}_N} L^{(h_{N,I})}(1, \chi) \ll (\log q)^{O(1)} (\log R)^{\mathcal{A}(V_N) - N - D},$$

where

$$D = d + \sum_{I \in \mathcal{S}^-(2k) \setminus \mathcal{I}_N} h_{N,I} + \sum_{\substack{I \in \mathcal{S}^-(2k) \cap \mathcal{I}_N \\ h_{N,I} \geq 1}} (h_{N,I} - 1)$$

and $\mathcal{A}(V_N)$ is defined in Section 7. Indeed, when $h_{N,I} = 0$ with $I \in \mathcal{S}^-(2k) \cap \mathcal{I}_N$, we use (12.4) to find that have that

$$L(1, \chi) \ll \frac{(\log q)^2}{\log Q} \leq \frac{(\log q)^2}{\log R}.$$

Otherwise, we use the bound $L^{(h_{N,I})}(1, \chi) \ll (\log q)^{h_{N,I}+1}$. Putting these estimates together yields (12.6).

Proof of Lemma 12.1. (a) Here $\gamma_{j_{N+1}} > 0$. We make the change of variables

$$s'_j = s_j \ (1 \leq j < j_{N+1}), \ s'_j = s_{j+1} \ (j_{N+1} \leq j \leq 2k - N), \ s'_{2k-N} = s_{j_{N+1}},$$

and similarly for the forms x_j and the parameters λ_j .

Next, we shift the s'_{2k-N} contour to the line $\operatorname{Re}(s'_{2k-N}) = -\epsilon$ for a small enough $\epsilon > 0$. The integral on the new contour is $\ll (\log R)^{O(1)}/T$, which is negligible, and we are left with having to analyze the pole contributions. The poles occur when $L_{N, I_{N+1}}(s') = 0$ for some $I_{N+1} \in \mathcal{S}^+(2k) \setminus \mathcal{I}_N$ such that the coefficient of s'_{2k-N} in $L_{N, I_{N+1}}$ is non-zero. As we discussed in the previous section, imposing the relation $L_{N, I_{N+1}}(x') = 0$ to write the form x'_{2k-N} as a linear combination of x'_1, \dots, x'_{2k-N-1} , say $x'_{2k-N} = C(x'_1, \dots, x'_{2k-N-1})$. In particular, $L_{N, I}(x')$ becomes a linear form $L_{N+1, I}$ in the variables x'_1, \dots, x'_{2k-N-1} . We also set $E_{N+1} = L_{N+1, [2k]}$ and let \mathcal{I}_{N+1} be the set of $I \subset [2k]$ such that $L_{N+1, I} = 0$.

The generic order of the pole at $s'_{2k-N} = C(s'_1, \dots, s'_{2k-N-1})$ is

$$(12.7) \quad m = \sum_{I \in (\mathcal{I}_{N+1} \setminus \mathcal{I}_N) \cap \mathcal{S}^+(2k)} (h_{N, I} + 1) - \nu,$$

where ν is the generic order of the zero of

$$G(s) \prod_{I \in \mathcal{S}^-(2k) \cap (I_{N+1} \setminus I_N)} L^{(h_{N, I})}(1 + L_{N, I}(s), \chi)$$

at the same point. In particular, $\nu = 0$ if $d = 0$ (so that $G(s) = F(L_{N, \{1\}}(s), \dots, L_{N, \{2k\}}(s))$) and $h_{N, I} = 0$ for all $I \in \mathcal{S}^-(2k) \cap (\mathcal{I}_{N+1} \setminus \mathcal{I}_N)$. By a direct computation, we then find that

$$(12.8) \quad X_{2k-1} + m = \#(\mathcal{I}_{N+1} \cap \mathcal{S}^+(2k)) - \sum_{I \in \mathcal{S}^-(2k) \cup (\mathcal{S}^+(2k) \setminus \mathcal{I}_{N+1})} h_{N, I} - \nu.$$

We note that $m \geq 1$ for all $N \leq 2k - 1$ when $k \geq 2$ and $\nu = 0$; otherwise, we would have that $\mathcal{S}^+(2k) \subset \mathcal{I}_N$, which is impossible because the dimension of V_N is N , whereas the dimension of the span of the linear forms s_I , $I \in \mathcal{S}^+(2k)$, is $2k$.

We want to understand the pole contribution when $m \geq 1$. We separate two subcases:

Case 1 of the proof of Lemma 12.1: $N = 2k - 1$. In this case, we have that $s'_j = L_{2k-1, \{j\}}(s'_1) = a_j s'_1$ for all j , where $a_j \in \mathbb{Q}$. Then the pole occurs necessarily when $s'_1 = 0$. Thus $\mathcal{I}_{2k} = \mathcal{S}(2k)$, and we obtain an evaluation of $J(R)$ as finite linear combination of powers of $\log R$, the highest of which has exponent

$$X_{2k-1} + m - 2k = 2^{2k-1} - 2k - 1 - \sum_{I \in \mathcal{S}^-(2k)} h_{2k-1, I} - \nu,$$

up to an admissible error. The coefficients of the polynomial in $\log R$ are given in terms of the derivatives $L^{(j)}(1, \chi)$. Specifically, the coefficient of $(\log R)^{X_{2k-1} + m - 2k - h}$, $0 \leq h \leq m - 1$, is a linear combination of products of the form

$$\left(\frac{\varphi(q)}{q} \right)^{M_{2k}} \prod_{I \in \mathcal{S}^-(2k)} L^{(h_{2k, I})}(1, \chi),$$

with the coefficients of this linear combination being $\ll 1$, and with the parameters $h_{2k, I}$ satisfying $h_{2k, I} \geq h_{2k-1, I}$ with equality if $I \in \mathcal{I}_{2k-1} \setminus \{0\}$, and $\sum_{I \in \mathcal{S}^-(2k)} (h_{2k, I} - h_{2k-1, I}) \leq h$. Arguing as

in the proof of (12.6), we find that

$$J(R) \ll \frac{(\log q)^{O(1)} (\log R)^{X_{2k-1}+m-2k-h}}{(\log Q)^{\#\{I \in \mathcal{S}^-(2k): h_{2k,I}=0\}}} \leq \frac{(\log q)^{O(1)} (\log R)^{2^{2k-1}-2k-1-\sum_{I \in \mathcal{S}^-(2k)} h_{2k,I}}}{(\log R)^{\#\{I \in \mathcal{S}^-(2k): h_{2k,I}=0\}}},$$

where we used that $Q \leq R$, $\nu \geq 0$ and $h \geq \sum_{I \in \mathcal{S}^-(2k)} (h_{2k,I} - h_{2k-1,I})$. We thus conclude that

$$J(R) \ll \frac{(\log q)^{O(1)}}{(\log R)^{2k+1}} \ll \frac{1}{\log R}.$$

This completes the proof of the lemma in this case (the linear combination is empty).

Case 2 of the proof of Lemma 12.1: $N \leq 2k - 2$. Arguing as in the proof of part (b) of Lemma 8.4, the contribution of the pole $s'_{2k-N} = C(s'_1, \dots, s'_{2k-N-1})$ to $J(R)$ can be seen to be a linear combination of terms of the form

$$\begin{aligned} & \left(\frac{\varphi(q)}{q} \right)^{M_{N+1}} \frac{(\log R)^{X_N+m-h-N-1-d}}{(2i\pi)^{2k-N-1}} \int \cdots \int_{\substack{\operatorname{Re}(s_j)=\lambda'_j/\log R, |\operatorname{Im}(s_j)| \leq T \\ 1 \leq j \leq 2k-N-1}} \tilde{G}(\mathbf{s}) R^{E_{N+1}(\mathbf{s})} \\ & \times \prod_{I \in \mathcal{S}^-(2k)} L^{(h_{N+1,I})} (1 + L_{N+1,I}(\mathbf{s}), \chi) \\ & \times \prod_{I \in \mathcal{S}^+(2k) \setminus \mathcal{I}_{N+1}} \zeta_q^{(h_{N+1,I})} (1 + L_{N+1,I}(\mathbf{s})) ds_{2k-N-1} \cdots ds_1 \end{aligned}$$

plus an error term of size $O(1/\log R)$, where $h \in \{0, \dots, m-1\}$, $h_{N+1,I} \geq h_{N,I}$ with equality if $I \in \mathcal{I}_{N+1} \setminus \{0\}$, and $\sum_{I \in \mathcal{S}^-(2k) \cup (\mathcal{S}^+(2k) \setminus \mathcal{I}_{N+1})} (h_{N+1,I} - h_{N,I}) \leq h$. Relation (12.8) then implies that the power of $\log R$ is then $X_{N+1} - N - 1 - d'$ with

$$d' = d + \nu + h - \sum_{I \in \mathcal{S}^-(2k) \cup (\mathcal{S}^+(2k) \setminus \mathcal{I}_{N+1})} (h_{N+1,I} - h_{N,I}) \geq 0.$$

Moreover, $X_{N+1} - d' = X_N + m - h \geq N + 1$. This completes the proof of part (a).

(b) Here $\gamma_{j_{N+1}} < 0$. We treat this case using the same argument as in part (b) of Lemma 8.4, with the difference that the contours of $s_{2k-N}, s_{2k-N-1}, \dots, s_{j_{N+1}}$ are shifted to the lines $\operatorname{Re}(s_j) = \lambda^j / ((\log q) + (\log T)^{3/2})$, $j_{N+1} \leq j \leq 2k - N$. Since $q \leq e^{\sqrt{\log R}}$ by assumption, we find that

$$J(R) \ll T^{-1+o(1)},$$

as needed. □

Proof of Lemma 12.2. We separate three cases.

Case 1 of the proof of Lemma 12.2: $N = 2k$. Here $\mathcal{I}_{2k} = \mathcal{S}(2k)$ and thus $J(R) \ll (\log q)^{O(1)} / (\log R)^{2k+1}$ by (12.6), which proves Lemma 12.2 in this case.

Case 2 of the proof of Lemma 12.2: $N = 2k - 1$. As in Case 1 of the proof of Lemma 12.1, we have that $s_j = L_{2k-1, \{j\}}(s_1) = a_j s_1$ for all j , where $a_j \in \mathbb{Q}$. Then

$$\begin{aligned} J(R) &= \left(\frac{\varphi(q)}{q} \right)^{M_{2k-1}} \frac{(\log R)^{X_{2k-1}-2k+1-d}}{2\pi i} \prod_{I \in \mathcal{S}^-(2k) \cap \mathcal{I}_{2k-1}} L^{(h_{2k-1, I})}(1, \chi) \\ &\quad \times \int_{\substack{\operatorname{Re}(s_1) = \lambda_1 / \log R \\ |\operatorname{Im}(s_1)| \leq T}} G_{2k-1}(s_1) \prod_{I \in \mathcal{S}^-(2k) \setminus \mathcal{I}_{2k-1}} L^{(h_{2k-1, I})}(1 + a_I s_1, \chi) \\ &\quad \times \prod_{I \in \mathcal{S}^+(2k) \setminus \mathcal{I}_{2k-1}} \zeta_q^{(h_{2k-1, I})}(1 + a_I s_1) ds_1, \end{aligned}$$

We first show a crude bound on $J(R)$, that will allow us to focus on a more convenient subcase. We move the line of integration to $\operatorname{Re}(s_1) = \lambda / \log(qT)$ and use (12.3) to find that the integral over s_1 is $\ll (\log(q \log R))^{O(1)}$. Together with (12.6) and Proposition 7.1, this implies that

$$J(R) \ll (\log(q \log R))^{O(1)} \cdot (\log R)^{\binom{2k}{k} - 2k - 1},$$

unless $d = 0$, $h_{2k-1, I} \in \{0, 1\}$ for all $I \in \mathcal{S}^-(2k) \cap \mathcal{I}_{2k-1}$, $h_{2k-1, I} = 0$ for all $I \in \mathcal{S}(2k) \setminus \mathcal{I}_{2k-1}$, half of the a_j 's are $+b$ and the other half are $-b$, for some $b \neq 0$.

We have thus reduced proving the lemma to the case when $d = 0$, $h_{2k-1, I} \in \{0, 1\}$ for all $I \in \mathcal{S}^-(2k) \cap \mathcal{I}_{2k-1}$, $h_{2k-1, I} = 0$ for all $I \in \mathcal{S}(2k) \setminus \mathcal{I}_{2k-1}$, half of the a_j 's are $+b$ and the other half are $-b$, where $b \neq 0$. In particular, we find that $\mathcal{I}_{2k-1} \cap \mathcal{S}^-(2k) = \emptyset$ and that $\#(\mathcal{I}_{2k-1} \cap \mathcal{S}^+(2k)) = \binom{2k}{k} - 1$, so that

$$\begin{aligned} J(R) &= \left(\frac{\varphi(q)}{q} \right)^{M_{2k-1}} \frac{(\log R)^{\binom{2k}{k} - 2k}}{2\pi i} \int_{\substack{\operatorname{Re}(s_1) = \lambda / \log R \\ |\operatorname{Im}(s_1)| \leq T}} G(s_1) \prod_{I \in \mathcal{S}^-(2k)} L(1 + a_I s_1, \chi) \\ &\quad \times \prod_{I \in \mathcal{S}^+(2k) \setminus \mathcal{I}_{2k-1}} \zeta_q(1 + a_I s_1) ds_1. \end{aligned}$$

Since $\nu = 0$ here, we saw before that the integrand has a genuine pole of order $m \geq 1$ at $s_1 = 0$ by a dimension argument. In fact, we have that $m \geq 2$: indeed, we know that $[2k] \in \mathcal{I}_{2k-1}$ by our assumption that $E_{2k-1} = 0$, so that $I \in \mathcal{I}_{2k-1}$ if and only if $[2k] \setminus I \in \mathcal{I}_{2k-1}$. In particular, since we know that there is at least one $I \in \mathcal{S}^+(2k) \setminus \mathcal{I}_{2k-1}$, there must be at least two such I 's.

The presence of this pole make the estimation of $J(R)$ tricky. In particular, we cannot shift the contour to the line $\operatorname{Re}(s_1) = 0$ as in Case 2 of Section 8.2. Instead, we write

$$L(1 + a_I s_1, \chi) = L(\beta + a_I s_1, \chi) + \Delta(a_I s_1),$$

where

$$\Delta(s) := L(1 + s, \chi) - L(\beta + s, \chi).$$

We thus find that

$$\Delta(s) = \int_{\beta}^1 L'(u + s, \chi) du \ll \frac{\log(q + |t|)}{\log Q} \quad (\sigma \geq -1/\log(q + |t|)),$$

by (12.3) and the assumption that $\beta > 1 - 1/(100 \log q)$. With this notation,

$$J(R) = M + E,$$

where

$$M = \left(\frac{\varphi(q)}{q} \right)^{M_{2k-1}} \frac{(\log R)^{\binom{2k}{k}-2k}}{2\pi i} \int_{\substack{\operatorname{Re}(s_1)=\lambda/\log R \\ |\operatorname{Im}(s_1)| \leq T}} G_{2k-1}(s_1) \prod_{I \in \mathcal{S}^-(2k)} L(\beta + a_I s_1, \chi) \\ \times \prod_{I \in \mathcal{S}^+(2k) \setminus \mathcal{I}_{2k-1}} \zeta_q(1 + a_I s_1) ds_1,$$

and E is a sum of similar expressions where at least one of the $L(\beta + a_I s_1, \chi)$ factors is replaced by $\Delta(a_I s_1)$.

First, we bound E . Moving s_1 to the line $\operatorname{Re}(s_1) = 1/\log(qT)$, we find that

$$E \ll \frac{(\log(q \log R))^{O(1)} (\log R)^{\binom{2k}{k}-2k}}{\log Q} \leq (\log(q \log R))^{O(1)} (\log R)^{\binom{2k}{k}-2k-1}.$$

Finally, we estimate M by moving the line of integration of s_1 to $\operatorname{Re}(s_1) = 0$, and use the fact that $G(s_1) \ll 1/(1 + |s_1|)^{2k}$ (note that the integrand is now analytic, since the pole of the ζ_q 's is annihilated by the zeroes of the $L(\cdot, \chi)$'s) to find that

$$J(R) = \left(\frac{\varphi(q)}{q} \right)^{M_{2k-1}} \frac{(\log R)^{\binom{2k}{k}-2k}}{2\pi} \int_{-T}^T G(it) \prod_{I \in \mathcal{S}^-(2k)} L(\beta + ia_I t, \chi) \\ \times \prod_{I \in \mathcal{S}^+(2k) \setminus \mathcal{I}_{2k-1}} \zeta_q(1 + ia_I t) dt + O((\log(q \log R))^{O(1)} (\log R)^{\binom{2k}{k}-2k-1}).$$

Note that $G(s) = F(a_1 s, \dots, a_{2k} s)$ here, because $d = 0$. When $|t| \leq \log R$, we use (6.6) to replace \widehat{h}_R by R^s/s , and when $|t| > \log R$ we use the bound $\widehat{h}_R(s) \ll R^\sigma/|s|$ by (6.5). Taking C to be large enough, we thus conclude that

$$\frac{J(R)}{(\log R)^{\binom{2k}{k}-2k}} = \frac{(\varphi(q)/q)^{M_{2k-1}}}{2\pi} \int_{-\infty}^{\infty} F(ia_1 t, \dots, ia_{2k} t) \prod_{I \in \mathcal{S}^-(2k)} L(\beta + ia_I t, \chi) \\ \times \prod_{I \in \mathcal{S}^+(2k) \setminus \mathcal{I}_{2k-1}} \zeta_q(1 + ia_I t) \prod_{j=1}^{2k} \frac{1 - 2^{-ia_j t}}{ia_j t} dt + O\left(\frac{(\log(q \log R))^{O(1)}}{\log R}\right).$$

This completes the proof of Lemma 12.2 in this case.

Case 3 of the proof of Lemma 12.2: $N \leq 2k - 2$. As in the corresponding case of the proof of Lemma 8.5, and using (12.6), we find that

$$J(R) \ll (\log(q \log R))^{O(1)} (\log R)^{X_N - N - d} \prod_{I \in \mathcal{S}^-(2k) \cap \mathcal{I}_N} |L^{(h_{N,I})}(1, \chi)| \\ \ll (\log(q \log R))^{O(1)} (\log R)^{\mathcal{A}(V_N) - N}.$$

Proposition 7.1(c) then implies that $\mathcal{A}(V_N) - N \leq \binom{2k}{k} - 2k - 2$, thus completing the proof of Lemma 12.2. \square

12.5. Lower bounds. In order to complete the proof of Theorem 1.5, we show that constant $c_k(\chi)$ in (12.5) is $\gg (\log q)^{-O(1)}$. In order to do so, we follow the argument of Section 9 and prove that, for any $\epsilon > 0$, there is a constant $c_k > 0$ such that

$$(12.9) \quad \mathcal{X}_{2k}(R) \geq c_k \frac{(\log R)^{\binom{2k}{k}-2k-\epsilon}}{(\log q)^{O(1)}} - O_\epsilon \left((\log(q \log R))^{O(1)} (\log R)^{\binom{2k}{k}-2k-1} \right),$$

provided that $\sqrt{\log Q} \geq \log R \geq 2c_1 \log q$, where c_1 is the constant appearing in (12.3). (For this section, all constants will be independent of ϵ , unless specified by a subscript, as above.)

We set

$$y = \exp\{(\log R)^{1-\epsilon'}\} \quad \text{and} \quad Y = \exp\{(\log R)^{1-\epsilon'/2}\},$$

where ϵ' will be taken to be small enough in terms of ϵ , and focus our attention on integers of the form $n = ap_1 \cdots p_k$ with $P^+(a) \leq y$, $a \leq Y$, and p_1, \dots, p_k are distinct primes such that $p_\ell > \sqrt{R}$ and $\chi(p_\ell) = -1$ for all $\ell \in \{1, \dots, k\}$. Then

$$\mathcal{X}_{2k}(R) \geq \frac{\prod_{p \leq R} (1 - 1/p)}{k!} \sum_{\substack{P^+(a) \leq y \\ a \leq Y}} \sum_{\substack{p_j > \sqrt{R} \\ \chi(p_j) = -1 \\ 1 \leq j \leq k}} \frac{\mu^2(p_1 \cdots p_k)}{ap_1 \cdots p_k} \left(\sum_{\ell=1}^k \sum_{\substack{R/(2p_\ell) < d \leq R/p_\ell \\ d|a}} \chi(d) \right)^{2k}.$$

The next step is to drop the condition that $a \leq Y$ by an application of Rankin's trick and to remove the condition that the p_j 's are distinct. We further replace the sharp cut-off $R/(2p_\ell) < d \leq R/p_\ell$ by the smooth cut-off $h(\frac{\log(dp_\ell)}{\log R}) - h(\frac{\log(2dp_\ell)}{\log R})$, where $h(x) = 1$ for $x \leq 1 - 1/(\log R)^B$ and $h(x) = 0$ for $x \geq 1$, with B is sufficiently large. To conclude, we have that

$$\begin{aligned} \mathcal{X}_{2k}(R) &\geq \frac{\prod_{p \leq R} (1 - 1/p)}{k! (\log R)^k} \sum_{\substack{P^+(a) \leq y \\ \chi(p_j) = -1 \\ 1 \leq j \leq k}} \sum_{\substack{\sqrt{R} < p_j \leq R \\ 1 \leq j \leq k}} \frac{\prod_{j=1}^k \log p_j}{ap_1 \cdots p_k} \left(\sum_{\ell=1}^k \sum_{d|a} \chi(d) w\left(\frac{\log(dp_\ell)}{\log R}\right) \right)^{2k} \\ &\quad + O\left(\frac{1}{\log R}\right), \end{aligned}$$

where $w(x) = h(x) - h(x + \log 2 / \log R)$.

The rest of the proof follows the argument of Section 9, with a small twist, as we will explain in the end. We expand the $2k$ -th power and focus on a convenient subset of summands. We then conclude that

$$\mathcal{X}_{2k}(R) \geq \frac{\prod_{y < p \leq R} (1 - 1/p)}{k!} \sum_{\mathbf{J} \in \mathcal{J}} X(\mathbf{J}) + O\left(\frac{1}{\log R}\right)$$

with

$$X(\mathbf{J}) = \frac{1}{(\log R)^k} \sum_{\substack{P^+(d_j) \leq y \\ 1 \leq j \leq 2k}} \frac{\chi(d_1) \cdots \chi(d_{2k})}{[d_1, \dots, d_{2k}]} \prod_{\ell=1}^k \sum_{\sqrt{R} < p_\ell \leq R} \frac{(1 - \chi(p_\ell)) \log p_\ell}{2p_\ell} \prod_{j \in J_\ell} w\left(\frac{\log(p_\ell d_j)}{\log R}\right),$$

where \mathbf{J} is as in Section 9. We set

$$S := \sum_{\sqrt{R} < p \leq R} \frac{(1 - \chi(p)) \log p}{2p} \asymp \log R$$

and write \mathcal{L} for the set of $\ell \in \{1, \dots, k\}$ such that $J_\ell \neq \emptyset$. Then, using Perron's formula $2k$ times to write each appearance of w as an integral of \widehat{w}_R , we find that

$$X(\mathbf{J}) = \frac{S^{k-\#\mathcal{L}}(\log R)^{-k}}{(2\pi i)^{2k}} \int \cdots \int_{\substack{\operatorname{Re}(s_j)=1/\log R \\ 1 \leq j \leq 2k}} \sum_{\substack{P^+(d_j) \leq y \\ 1 \leq j \leq 2k}} \frac{\prod_{j=1}^{2k} \chi(d_j) d_j^{-s_j}}{[d_1, \dots, d_{2k}]} \left(\prod_{\ell \in \mathcal{L}} \sum_{p_\ell > \sqrt{R}} \frac{(1 - \chi(p_\ell)) \log p_\ell}{2p_\ell^{1+s_{J_\ell}}} \right) \\ \times \left(\prod_{j=1}^{2k} \widehat{w}_R(s_j) \right) ds_1 \cdots ds_{2k} + O\left(\frac{1}{\log R}\right),$$

where $s_J = \sum_{j \in J} s_j$, as usual, and the condition that $p_\ell \leq R$ was dropped because it is encoded in the support of w . By possibly re-indexing the variables s_1, \dots, s_{2k} , we may assume that $\mathcal{L} = \{1, \dots, L\}$, where $L = \#\mathcal{L}$, and that $\max J_\ell = 2k - L + \ell$ for all $\ell \in \{1, \dots, L\}$ with $L = \#\mathcal{L}$. As in Section 9, we will move the variables $s_{2k-L+1}, \dots, s_{2k}$ to the left. We note that

$$\sum_{p > \sqrt{R}} \frac{(1 - \chi(p)) \log p}{p^{1+s}} = -\frac{\zeta'}{\zeta}(1+s) + \frac{L'}{L}(1+s, \chi) + O(1) - \sum_{p \leq R^{1/2}} \frac{\log p}{p^{1+s}}$$

for $\operatorname{Re}(s) \geq -1/3$. The above has simple poles at $s = 0$ and $s = \beta - 1$, each of residue 1. Therefore, using the argument leading to (9.12), we find that

$$(12.10) \quad X(\mathbf{J}) = \sum_{\substack{\epsilon_\ell \in \{0, \beta-1\} \\ 1 \leq \ell \leq L}} \frac{S^{k-L}(\log R)^{-k}}{2^L (2\pi i)^{2k-L}} \int \cdots \int_{\substack{\operatorname{Re}(s_j)=1/\log R \ (1 \leq j \leq 2k-L) \\ s_{J_\ell} = \epsilon_\ell \ (1 \leq \ell \leq L)}} \sum_{\substack{P^+(d_j) \leq y \\ 1 \leq j \leq 2k}} \frac{\prod_{j=1}^{2k} \mu(d_j) d_j^{-s_j}}{[d_1, \dots, d_{2k}]} \\ \times \left(\prod_{j=1}^{2k} \widehat{w}_R(s_j) \right) ds_1 \cdots ds_{2k-L} + O\left(\frac{1}{\log R}\right).$$

The above expression is sufficient for handling the terms $\mathbf{J} \in \mathcal{J}$ with at least one J_ℓ of odd cardinality: following the argument of Section 9.5 with the obvious modifications implies that

$$(12.11) \quad X(\mathbf{J}) \ll (\log(q \log R))^{O(1)} \frac{(\log y)^{\binom{2k}{k} - 2k - 1 + L}}{(\log R)^L},$$

where we used the fact that $\sup_{q < p \leq Q} (1 + \chi(p))/p \ll 1$.

However, we need to be more careful on our lower bound for the main term, that is to say for $X(\mathbf{J})$ with $\#J_\ell = 2$ for all ℓ . First of all, by relabelling our variables, we may assume that $J_\ell = \{\ell, \ell + k\}$ for all ℓ . In our expression (12.10) for $X(\mathbf{J})$, we see that $s_{J_\ell} = \epsilon_\ell \in \{0, \beta - 1\}$ implies that $s_{\ell+k} = -s_\ell + O(1/\log Q)$. We want to replace $s_{\ell+k}$ by $-s_\ell$. This introduces an error that we will control by an application of the mean value theorem. In particular, we need to understand the derivative of the integrand. If

$$G(\mathbf{s}) = \sum_{\substack{P^+(d_j) \leq y \\ 1 \leq j \leq k}} \frac{\prod_{j=1}^{2k} \chi(d_j) d_j^{-s_j}}{[d_1, \dots, d_{2k}]},$$

then $-G'(\mathbf{s})/G(\mathbf{s})$ equals $\sum_{I \in S^*(2k)} \sum_{p \leq y} \chi^{\#I}(p) \log p / p^{1+s_I}$, plus lower order terms, so that $G'(\mathbf{s}) \ll (\log y) G(\mathbf{s})$ for the vectors \mathbf{s} we are considering. Similarly, $\widehat{w}_R(s+\epsilon)/R^{s+\epsilon} = \widehat{w}_R(s)/R^s +$

$O(\epsilon/(|s| + 1))$ by (6.5). Since we also have that

$$\int \cdots \int_{\substack{\operatorname{Re}(s_j)=1/\log R \\ |s_{j+k}+s_j|\leq 1-\beta \\ 1\leq j\leq k}} |G(\mathbf{s})| \prod_{j=1}^k \frac{|ds_j|}{1+|s_j|^2} \ll (\log(q \log R))^{O(1)} \frac{(\log y)^{\binom{2k}{k}-k}}{(\log R)^k},$$

by the argument leading to (12.11), we conclude that

$$\begin{aligned} X(\mathbf{J}) &= \frac{(1 + R^{\beta-1})^k}{(2\pi i \log R)^k} \int \cdots \int_{\substack{\operatorname{Re}(s_j)=1/\log R \\ s_{j+k}=-s_j \\ 1\leq j\leq k}} G(\mathbf{s}) \left(\prod_{j=1}^{2k} \widehat{w}_R(s_j) \right) ds_1 \cdots ds_k \\ &\quad + O\left(\frac{\log y}{\log Q} \cdot \frac{(\log(q \log R))^{O(1)} (\log y)^{\binom{2k}{k}-k+1}}{(\log R)^k} \right). \end{aligned}$$

The main term can now be bounded from below as in Section 9.4. We thus arrive to the lower bound

$$X(\mathbf{J}) \geq c_k \frac{(\log y)^{\binom{2k}{k}-k} (\log R)^{-k}}{(\log q)^{O(1)}} - O\left(\frac{(\log(q \log R))^{O(1)} (\log y)^{\binom{2k}{k}-k+1}}{(\log R)^{k+2}} \right),$$

using that $\sum_{q < p \leq Q} (1 + \chi(p))/p \ll 1$ and $y \leq R \leq e^{\sqrt{\log Q}}$ here. This completes the proof of (12.9), and thus of Theorem 1.5(c).

REFERENCES

- [1] M. Balazard, M. Naimi and Y.-F. S. Pétermann, *Étude d'une somme arithmétique multiple liée à la fonction de Möbius*, Acta Arith. **132** (2008), 245–298.
- [2] R. de la Bretèche, *Estimation de sommes multiples de fonctions arithmétiques*, Compos. Math. **128** (2001), 261–298.
- [3] Harold Davenport, *Multiplicative number theory*, volume 74 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, third edition, 2000. Revised and with a preface by Hugh L. Montgomery.
- [4] F. Dress, H. Iwaniec and G. Tenenbaum, *Sur une somme liée à la fonction de Möbius*, J. reine angew. Math. **340** (1983), 53–58.
- [5] S. Eberhard, K. Ford and B. Green, *Permutations fixing a k -set*, IMRN, to appear. arXiv:1507.04465
- [6] P. Erdős, and R. R. Hall, *On the Möbius function*, J. Reine Angew. Math. **315** (1980), 121–126.
- [7] K. Ford, *The distribution of integers with a divisor in a given interval*, Ann. of Math. **168**, (2008), 367–433.
- [8] ———, *Rough integers with a divisor in a given interval*. J. Australian Math. Soc., to appear, 17 pages, arXiv:1901.02548.
- [9] J. Friedlander, *Sifting short intervals. II*, Math. Proc. Cambridge Philos. Soc., **92** (1982), 381–384.
- [10] J. Friedlander and H. Iwaniec, *Opera de cribro*. American Mathematical Society Colloquium Publications, 57. American Mathematical Society, Providence, RI, 2010.
- [11] A. Granville, *Cycle lengths in a permutation are typically Poisson*. Electron. J. Combin. **13** (2006), no. 1, Research Paper 107, 23 pp.
- [12] R. R. Hall and G. Tenenbaum, *Divisors*. Cambridge Tracts in Mathematics, 90. Cambridge University Press, Cambridge, 1988.
- [13] A. Goldston, J. Pintz and C. Y. Yıldırım, *Primes in tuples. I*. Ann. of Math. (2) **170** (2009), no.2, 819–862.
- [14] H. Iwaniec and E. Kowalski, *Analytic number theory*. American Mathematical Society Colloquium Publications, 53. American Mathematical Society, Providence, RI, 2004.
- [15] D. Koukoulopoulos, *Localized factorizations of integers*. Proc. Lond. Math. Soc. (3) **101** (2010), no. 2, 392–426.
- [16] ———, *Pretentious multiplicative functions and the prime number theorem for arithmetic progressions*. Compos. Math., **149** (2013), no. 7, 1129–1149.

- [17] ———, *On multiplicative functions which are small on average*. Geom. Funct. Anal. 23 (2013), no. 5, 1569–1630.
- [18] J. Maynard, *Small gaps between primes*, Ann. of Math. 181 (2015), 383–413.
- [19] Y. Motohashi, *A multiple sum involving the Möbius function*, Publ. Inst. Math. (Beograd) (N.S.) **76(90)** (2004), 31–39.
- [20] M. Rosen, *Number theory in function fields*. Graduate Texts in Mathematics, 210. Springer-Verlag, New York, 2002.
- [21] A. Selberg, *On an elementary method in the theory of primes*, Norske Vid. Selsk. Forh., Trondhjem 19 (1947), 64–67.
- [22] A. Selberg, *Collected papers. II*, Springer Collected Works in Mathematics, Springer, Heidelberg, 2014.
- [23] P. Pollack, *Bounded gaps between primes with a given primitive root*, Algebra Number Theory 8 (2014), 1769–1786.
- [24] T. Tao, *Polymath8b: Bounded intervals with many primes, after Maynard*, Blog note. <https://terrytao.wordpress.com/2013/11/19/polymath8b-bounded-intervals-with-many-primes-after-maynard/>
- [25] G. Tenenbaum, *Introduction to analytic and probabilistic number theory*. Translated from the second French edition (1995) by C. B. Thomas. Cambridge Studies in Advanced Mathematics, 46. Cambridge University Press, Cambridge, 1995.
- [26] E. C. Titchmarsh, *The theory of the Riemann zeta-function*. Second edition. Edited and with a preface by D. R. Heath-Brown. The Clarendon Press, Oxford University Press, New York, 1986.
- [27] Y. Zhang, *Bounded gaps between primes*, Ann. of Math. (2) 179 (2014), no. 3, 1121–1174.

AG: DÉPARTEMENT DE MATHÉMATIQUES ET DE STATISTIQUE, UNIVERSITÉ DE MONTRÉAL, CP 6128 SUCC. CENTRE-VILLE, MONTRÉAL, QC H3C 3J7, CANADA; AND DEPARTMENT OF MATHEMATICS, UNIVERSITY COLLEGE LONDON, GOWER STREET, LONDON WC1E 6BT, ENGLAND.

Email address: andrew@dms.umontreal.ca

DK: DÉPARTEMENT DE MATHÉMATIQUES ET DE STATISTIQUE, UNIVERSITÉ DE MONTRÉAL, CP 6128 SUCC. CENTRE-VILLE, MONTRÉAL, QC H3C 3J7, CANADA

Email address: koukoulo@dms.umontreal.ca

JM: MATHEMATICAL INSTITUTE, RADCLIFFE OBSERVATORY QUARTER, WOODSTOCK ROAD, OXFORD OX2 6GG, ENGLAND

Email address: james.alexander.maynard@gmail.com