Points on algebraic curves over function fields, primes in arithmetic progressions : beyond Bombieri-Pila and Bombieri-Vinogradov theorems

Alisa Sedunova

NNT : 2017SACLS178

THÈSE DE DOCTORAT

de

L'Université Paris-Saclay

École Doctorale de Mathématiques de la Région Paris-Sud ED 142

Établissement d'inscription : Université Paris-Sud
Établissement d'accueil : Georg-August-Universität Göttingen

Spécialité de doctorat : Mathématiques fondamentales

Alisa SEDUNOVA

Points entiers sur les courbes algébriques sur les corps de fonctions, les nombres premiers dans les progressions arithmétiques: au-delà des théorèmes de Bombieri-Pila et de Bombieri-Vinogradov.

Date de soutenance : 27 Juin 2017

Après avis des rapporteurs : Marc HINDRY, Université Paris Diderot
Yuri BILU, Université de Bordeaux

Jury de soutenance:

Étienne FOUVRY, Université Paris Sud	Président du jury
Harald HELFGOTT, Université Paris Diderot	Directeur de thèse
Marc HINDRY, Université Paris Diderot	Rapporteur
Yuri BILU, Université de Bordeaux	Rapporteur
Régis DE LA BRETÈCHE, Université Paris Diderot	Examinateur
Olivier RAMARÉ, Aix-Marseille Université	Examinateur

université PARIS-SACLAY | École doctorale de mathématiques Hadamard (EDMH)

Titre : Points sur les courbes algébriques sur les corps de fonctions, les nombres premiers dans les progressions arithmétiques: au-delà des théorèmes de Bombieri-Pila et de Bombieri-Vinogradov.

Mots Clefs : Bombieri-Pila, Bombieri-Vinogradov, points entiers, courbes elliptiques.

Résumé : E. Bombieri et J. Pila ont introduit une méthode qui donne des bornes pour le nombre de points entiers qui appartiennent à un arc donné (sous quelques hypothèses). Dans la partie algébrique nous généralisons la méthode de Bombieri Pila pour le cas des corps de fonctions en une variable, de genre 0. Ensuite, nous appliquons le résultat pour calculer le nombre de courbes elliptiques qui sont dans la même classe d'isomorphisme avec leurs coefficients dans une petite boîte. Une fois que nous avons prouvé ça, la question naturelle est de savoir si nous pouvons l'améliorer dans certains cas particuliers. Nous allons étudier le cas des courbes elliptiques en utilisant la partie connue de la conjecture de Birch Swinnerton-Dyer, les propriétés des fonctions de hauteur et les empilements compacts. Après, dans une partie analytique nous donnons une version explicite du théorème de Bombieri Vinogradov. Ce théorème est un résultat important, qui concerne le terme d'erreur dans le théorème de Dirichlet sur les progressions arithmétiques, pris en moyenne sur les modules q variant jusqu'à Q. Notre but est d'améliorer les résultats existant de cette façon (voir [1] et [13]), donc nous pouvons réduire la puissance du facteur logarithmique en utilisant l'inégalité de grand crible et l'identité de Vaughan.

Title : Points on algebraic curves over function fields, primes in arithmetic progressions: beyond Bombieri-Pila and Bombieri-Vinogradov theorems

Keys words : Bombieri-Pila, Bombieri-Vinogradov, integral points, elliptic curves.

Abstract : E.Bombieri and J.Pila introduced a method to bound the number of integral points in a small given box (under some conditions). In algebraic part we generalise this method to the case of function fields of genus 0 in ove variable. Then we apply the result to count the number of elliptic curves falling in the same isomorphic class with coefficients lying in a small box. Once we are done the natural question is how to improve this bound for some particular families of curves. We study the case of elliptic curves and use the fact that the necessary part of Birch Swinnerton-Dyer conjecture holds over function fields. We also use the properties of height functions and results about sphere packing.
In analytic part we give an explicit version of Bombieri-Vinogradov theorem. This theorem is an important result that concerns the error term in Dirichlet's theorem in arithmetic progressions averaged over moduli q up to Q. We improve the existent result of such type given in [1] and [13]. We reduce the logarithmic power by using the large sieve inequality and Vaughan identity.

Contents

Acknowledgements

First of all, I would like to thank my supervisor, Harald Helfgott for his continuous support of my PhD studies and his encouragement of my work. Besides my advisor, I would like to thank the rest of the thesis committee, and in particular, the rapporteurs, Marc Hindry and Yuri Bilu, for their careful proofreading of this thesis and their useful comments.

I would like to thank in particular Igor Shparlinski for raising the question of Section 2 and helping me a great deal in the process of solving the problem. I am also grateful to Yongqiang Zhao for his remarks on Section 2.

I am grateful to Henryk Iwaniec for crucial advice concerning Section 4.5 and Olivier Ramaré for his guidance in the starting stage of Section 4.5 as well as for helping in improving the exposition of Sections 4.1-4.4. I would like to thank as well Gérald Tenenbaum for an advice concerning Section 5.

I would also like to express my gratitude to the University Paris Sud for sorting out all kinds of problems in the last 4 years. My special thanks go to Étienne Fouvry for being helpful at all the stages of the writing of the present thesis and letting me speak several times at the seminar in Paris Sud. I am also indebted to Régis de la Bretèche for being attentive to my research and inviting me to talk in IHP. I deeply thank Ilya Shkredov for his endless patience and rewarding discussions.

Lastly, I am grateful to Université Paris Sud and Georg-August-Universität Göttingen for allowing me to pursue my graduate studies, and, in particular, for providing me with funding. My graduate experience benefited greatly from the courses I took and the opportunities I had while being a part of these places.

1 Introduction (version française)

1.1 La partie algébrique

Dans [4] E. Bombieri et J. Pila ont prouvé que

Théorème (Bombieri-Pila [4]). *Soit C est une courbe algébrique irréductible de degré d, Γ est un sous-ensemble de C dans un carré de taille N. Alors le nombre des points entiers dans Γ est borné par*

$$c(d,\varepsilon)|N|^{\frac{1}{d}+\varepsilon}$$

pour chaque $\varepsilon > 0$. La constante $c(d,\varepsilon)$ ne dépend pas de Γ.

Il y a beaucoup d'analogues de ce résultat très connu. Par exemple, nous pouvons être intéressé à trouver une borne pour un certain nombre de solutions de $f(x,y) = 0(\mathrm{mod}\, p)$ avec $x \in I$, $y \in J$, où I et J sont des intervalles courts dans $\mathbb{Z}/p\mathbb{Z}$ (voir [6], [8]). Ces résultats sont les p-analogues d'un travaill par Bombieri-Pila. Ici nous supposons que les longueurs des intervalles I et J sont considérablement plus courts que p pour les techniques standards comme, par exemple, les bornes de Weil deviennent inopérantes.

Nous pouvons aller plus loin et chercher un analogue sur corps de fonctions. Maintenant nous travaillons avec le corps fini $\mathbb{F}_{q^n}$, que nous voyons comme $\mathbb{F}_q[T]/f(T)$, où f est un polynôme irréductible fixe de degré n et T est comme d'habitude une variable formelle. Ce point de vue était utilisé par J. Cilleruelo et I. Shparlinski dans [7] pour les estimations d'un nombre de solutions de congruences polynomiales modulo un nombre premier avec les variables appartenait à des intervalles courts. Les mêmes auteurs ont formule [7, Problème 9], qui est résolu ici.

Nous allons prouver un analogue de ce résultat pour le cas de corps de fonctions. Soit $X, Y \in \mathbb{F}_q[T]$, où $q = p^\alpha$, $\alpha \in \mathbb{N}$ avec p un nombre premier, T est une variable formelle, ce qui signifie que

$$X = X(T) = a_0 + a_1 T + \ldots + a_n T^n,$$
$$Y = Y(T) = b_0 + b_1 T + \ldots + b_m T^m,$$

où tous les $a_i, b_j \in \mathbb{F}_q$, $i = 0, \ldots, \deg X = n$, $j = 0, \ldots, \deg Y = m$. Définissons la norme d'un polynôme $X \in \mathbb{F}_q[T]$ comme $|X| = q^{\deg X}$.

Disons que "un intervalle" I dans $\mathbb{F}_q[T]$ est un ensemble de polynômes de la forme $X(T)+Y(T)$, où $X(T)$ est un polynôme fixe et $Y(T)$ est un polynôme de degré borné par un nombre entier donné. Le longeur de I est donnée par $|I| = q^{\max \deg Y}$. Soit $\mathcal{C}$ est une courbe algébrique irréductible de degré d sur $\mathbb{F}_q[T]$, qui est décrit par un equation $F(X,Y) = 0$, où $F(X,Y) \in (\mathbb{F}_q[T])[X,Y]$. Nous écrivons S pour un ensemble de points de $\mathcal{C}$, qui sont aussi dans I^2, où I est un intervalle dans $\mathbb{F}_q[T]$ (fixons $X = 0$ dans le définition de I).

Nous allons formuler notre premier résultat.

Théorème 1. *Soit $\mathcal{C}$ est une courbe algébrique irréductible de degré d sur $\mathbb{F}_q[T]$, $q = p^\alpha$, p est un nombre premier, $\alpha \in \mathbb{N}$. Écrivons S pour l'ensemble de points de $\mathcal{C}$ dans I^2 pour quelque intervalle I de taille $|I| = q^{n+1}$. Alors*

$$|S| \ll_{d,\varepsilon} |I|^{\frac{1}{d}+\varepsilon}.$$

Nous nous pouvons poser le question: pourquoi nous pouvons pas simplement suivre l'approche de Bombieri-Pila pour obtenir le Théorème 1? Malheureusement, dans ce cas-la nous allons avoir les problèmes avec un analogue de Lemme 2 de [4], parce que il n'y a pas un analogue nécessaire d'un valeur moyen dans le cas de corps de fonctions (voir [42, Lemme 1]). Il semble y avoir au moins deux façons plausibles pour éviter cette difficulté. La première consiste à obtenir une variante de corps de fonctions du Théorème 14 dans l'article de Heath-Brown [21]. La deuxième technique est d'adapter la méthode de Helfgott-Venkatesh [25]. Ici nous allons suivre la deuxième approche qui demande moins des calculs longs et techniques. Nous aurons besoin d'analogues de Propositions 3.1 et 3.2 de [25]. Le développement des idées originales de [4] avec une adaptation de certains résultats de [25] va donner le Théorème 1.

Après ça, nous allons utiliser le Théorème 1 pour obtenir certaines applications, telles que les bornes sur le nombre de classes d'isomorphismes qui sont représentés par les courbes elliptiques $E_{a,b}$ avec les coefficients $a, b \in \mathbb{F}_q[T]$, qui sont aussi dans une petite boîte, disons, I^2. En utilisant ce résultat nous pouvons estimer le nombre de courbes elliptiques situées dans une classe d'isomorphismes donné avec leur coefficients dans une petite boîte. Pour poursuivre, nous allons travailler avec des idées proposées dans [8].

Soit $q = p^\alpha$, où p est un nombre premier, $\alpha \in \mathbb{N}$. Considérons la famille de courbes elliptiques $E_{a,b}$

$$E_{a,b} : Y^2 = X^3 + aX + b,$$

où X et Y sont appartiennent de $\mathbb{F}_q[T]$ et a, b sont quelques coefficients dans $\mathbb{F}_q[T]$ avec la propriété que $4a^3 + 27b^2 \neq \mathbf{0}$. Soit $f \in \mathbb{F}_q[T]$ est un polynôme irreductible. Nous disons que les deux courbes $E_{a,b}$ et $E_{c,d}$ sont isomorphes sur $\mathbb{F}_q[T]/(f)$ si et seulement si il existe t inversible modulo f tel que

$$at^4 \equiv c (\mathrm{mod}\, f) \quad \text{and} \quad bt^6 \equiv d (\mathrm{mod}\, f).$$

L'existence d'un isomorphisme entre $E_{a,b}$ et $E_{c,d}$ implique que

$$a^3 d^2 \equiv c^3 b^2 (\mathrm{mod}\, f). \tag{1}$$

Soit $\lambda \in \mathbb{F}_q[T]$. Nous écrivons $N_\lambda(I^2)$ pour le nombre de solutions de la congruence

$$a^3 \equiv \lambda b^2 (\mathrm{mod}\, f), \quad (a, b) \in I^2.$$

Nous allons donner une borne supérieure de $N_\lambda(I^2)$, qui implique les bornes supérieures pour le nombre de courbes elliptiques $E_{a,b}$ avec les coefficients $a, b \in I$ que se trouvent dans les mêmes classes d'isomorphismes.

Maintenant nous pouvons donner une bonne borne pour $N_\lambda(I^2)$.

Théorème 2. *Soit I est une intervalle de polynômes de degré plus petit ou égal à d avec les coefficients dans $\mathbb{F}_q$ et $|I| = q^d$. Pour chaque polynôme irréductible $f \in \mathbb{F}_q[T]$ tel que $1 \leqslant |I| \leqslant |f|^{\frac{1}{9}}$ et pour chaque $\lambda \in \mathbb{F}_q[T]$ nous avons*

$$N_\lambda(I^2) \leqslant |I|^{\frac{1}{3}+o(1)}.$$

Soit K le corps de fonctions rationnelles sur une courbe algébrique du genre g sur le corps constant k de la caractéristique 0. Notons par M_K l'ensemble de toutes les valuations v de K. Pour un sous-ensemble fini $S \subset M_K$ nous écrivons $\mathcal{O}_S$ pour

l'anneau de points S-entiers de K. Considérons une courbe elliptique non constante E donnée par une équation Weierstrass minimale

$$y^2 + a_1xy + a_3y = x^2 + a_2x^2 + a_4x + a_6,$$

où tous les coefficients $a_1, a_2, a_3, a_4, a_6 \in \mathcal{O}_S$. L'ensemble de points S-entiers d'une courbe elliptique E est

$$E(\mathcal{O}_S) = \{P \in E(K) : x(P), y(P) \in \mathcal{O}_S\}.$$

Nous allons améliorer le résultat de Théorème 1 pour la courbe elliptique E du conducteur N. Notre deuxième résultat important est

Théorème 3. *Soit q est la puissance d'un nombre premier, $\mathbb{F}_q[T]$ est un corps de polynômes en variable formelle T avec les coefficients dans le corps fini $\mathbb{F}_q$ d'ordre p. Soit E est une courbe elliptique sur $\mathbb{F}_q[T]$ avec un conducteur N. Supposons que les points entiers de E sont sur un modèle minimal. Alors le nombre de points entiers appartiennent de E satisfait*

$$\#E(\mathbb{F}_q[T]) \leq \exp\left(c\frac{\deg N_E}{\log\deg N_E}\right),$$

où la constante implicite est absolue et N_E est un degré du conducteur N de la courbe E.

Les résultats utiles dans la théorie de courbes elliptiques

Nous allons commencer avec le petit rappel de résultats plus basiques sur les courbes elliptiques sur corps de fonctions. Généralement, les outils qui nous permettent de procéder sont que la partie nécessaire de la conjecture célèbre de Birch et Swinnerton-Dyer est vrai dans le cas du corps de fonctions, ainsi que les limites pour le rang analytique sur le corps de fonctions sont connues, grâce à la formule explicite donnée par Brumer dans [5]. Nous utilisons également la technique de Helfgott (voir [24], basé sur des idées de Silverman [40] et [39]) pour obtenir une borne supérieure pour le nombre de points entiers sur E en termes de son rang algébrique. Cependant, cela nous amène à des résultats qui dépendent de la courbe C. Pour se débarrasser de cette dépendance, nous devons travailler avec l'estimation de la sorte $\#E(\mathbb{F}_q[T]) \ll c^{\operatorname{rank} E}$ plus attentivement. A savoir, nous étendons la méthode développée par Helfgott-Venkatesh dans [23]. Nous optimisons la taille de c en appliquant les résultats d'empilement de sphères [27] (voir aussi Lemme 3).

Rapellons qu'une courbe elliptique sur $K = \mathbb{F}_q[T]$ est une courbe lisse, projective, algébrique absolument irréductible de genre 1 sur K avec un K point-rationnel $\mathcal{O}$, qui marche comme un élément d'identité dans le groupe $E(K)$ des K, qui consiste en points rationnels appartiennent de E. Ce groupe est aussi appelé le groupe de Mordell-Weil de E. Il y a quelques définitions équivalents de la courbe elliptique E/K. Nous allons utiliser ici le fait que chaque courbe elliptique E/K peut être écrit sous la equation de Weierstrass (longue)

$$Y^2Z + a_1XYZ + a_3YZ^2 = X^3 + a_2X^2Z + a_4XZ^2 + a_6Z^3,$$

où $a_1, a_3, a_2, a_4, a_6 \in \mathbb{F}_q$. Si le caractéristique du corps n'est pas égal à $2, 3$, nous pouvons simplifier l'équation ci-dessus par deux changements de variables consécutifs

- $(X', Y', Z') = \left(X, Y + \frac{a_1}{2}X, Z\right)$ (celui-ci nous permet de disparaître le terme XYZ).
- $(X', Y', Z') = \left(X + \frac{a_2}{3}, Y + \frac{a_3}{2}X, Z\right)$ (celui-ci nous permet de disparaître les termes X^2 et Y).

Nous pouvons formuler ces résultats comme le théorème suivant.

Théorème. *Soit E est une courbe elliptique sur le corps K. Alors E est isomorphe à la courbe qui est décrit par*

$$Y^2Z + a_1XYZ + a_3YZ^2 = X^3 + a_2X^2Z + a_4XZ^2 + a_6Z^3,$$

où $a_1, a_3, a_2, a_4, a_6 \in K$. Si en plus $\operatorname{char} K \neq 2, 3$, *donc E est isomorphe à la courbe qui est décrit par*

$$Y^2Z = X^3 + dXZ^2 + bZ^3,\ a, b \in K.$$

Pour définir le discriminant $\Delta(E)$ et j-invariant de E nous avons besoin de quelques notations standards

$$\begin{aligned}
b_2 &= a_1^2 + 4a_2,\\
b_4 &= a_1a_3 + 2a_4,\\
b_6 &= a_3^2 + 4a_6,\\
b_8 &= b_2a_6 - a_1a_3a_4 + a_2a_3^2 - a_4^2,\\
c_4 &= b_2^2 - 24b_4,\\
c_6 &= -b_2^3 + 36b_2b_4 - 216b_6.
\end{aligned}$$

Alors le discriminant $\Delta(E)$ et j-invariant de la courbe elliptique E sont

$$\Delta(E) = -b_2^2b_8 - 8b_4^3 - 27b_6^2 + 9b_2b_4b_6, \quad j(E) = \frac{c_4}{\Delta(E)}.$$

Le condition de la texture lisse est équivalent àu fait que $\Delta(E) \neq 0$. Pour la simplicité nous supposons maintenant que $\operatorname{char} K \neq 2, 3$. Écrivons $x(P)$ et $y(P)$ pour les fonctions de coordonnes du point P. Pour les points $P_1 = (x_1, y_1), P_2 = (x_2, y_2)$ avec les coordonnes $x_1, x_2, y_1, y_2 \in K$ nous pouvons définir la somme $P_1 + P_2$ et le double $2P$.

- $P_2 = -P_1 : \quad P_1 + P_2 = \mathcal{O}$.
- $P_2 = P_1,\ y_1 = y_2 = 0 : \quad P_1 + P_2 = 2P = \mathcal{O}$.
- $P_2 = P_1 = (x, y),\ y \neq 0$. Alors la formule de la duplication est donne par

$$\begin{aligned}
x(P_1 + P_2) &= \frac{(3x + a)^2 - 8xy^2}{4y^2},\\
y(P_1 + P_2) &= \frac{x^6 + 5ax^4 + 20bx^3 - 5a^2x^2 - 4abx - a^3 - 8b^2}{8y^3}.
\end{aligned}$$

- $P_2 \neq \pm P_1$. Alors la formule de la sommation est donne par

$$\begin{aligned}
x(P_1 + P_2) &= \frac{x_1x_2^2 + x_1^2x_2 - 2y_1y_2 + a(x_1 + x_2) + 2b}{(x_1 - x_2)^2},\\
y(P_1 + P_2) &= \frac{W_2y_2 - W_1y_2}{(x_2 - x_2)^3},
\end{aligned}$$

où

$$W_1 = 3x_1x_2^2 + x_2^3 + a(x_1 + 3x_2) + 4b,$$
$$W_2 = 3x_1^2x_2 + x_1^3 + a(3x_1 + x_2) + 4b.$$

Nous pouvons distinguer deux type de points $P \in E(K)$. Disons que le point P est d'un ordre fini si et seulement si il y a un nombre $n \in \mathbb{N}$ tel que

$$nP = \underbrace{P + \ldots + P}_{n \text{ fois}} = \mathcal{O}.$$

Sinon, disons que P est un point d'un ordre infini.

Le résultat très important de Mordell-Weil dit que sous cette addition l'ensemble de points de la courbe E avec les coordonnes appartiennent de K est vraiment un groupe qui est en plus de type fini. Le résultat de Mordell-Weil classique a été généralisé au cas des corps de fonctions par Lang et Néron, où $E(K)$ est un groupe abélien fini. En conséquence de ce résultat, le groupe de torsion $E(K)_{\text{tors}}$ (le groupe de K-points sur E d'ordre fini) est fini et isomorphe à un groupe de forme

$$\mathbb{Z}/m\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z},$$

où m divise n et p ne divise pas m. Définissons un algébrique $\text{rank}(E)$ d'une courbe elliptique E/K comme le nombre de points indépendants d'ordre infini dans $E(K)$, ce qui revient le nombre de copies de $\mathbb{Z}$ en $E(K)$.

Soit v est une classe d'équivalence de valuations sur K. Rappellons que la valuation sur un corps K est une généralisation de la norme p-adique. En fait, c'est une fonction $|\cdot|_v : K \to \mathbb{R}$ avec les propriétés suivantes: pour touts $x, y \in K$ nous avons

- $|x|_v \geq 0$, $|x| = 0$ si et seulement si $x = 0$;
- $|xy|_v = |x|_v \cdot |y|_v$;
- si $|x|_v \leq 1$, alors $|1 + x|_v \leq C$ pour quelque constante $C \geq 1$ qui dépend pas de x.

Notons que si la valuation $|\cdot|_v$ satisfait la dernière condition avec $C = 2$, alors elle satisfait aussi l'inégalité triangulaire

$$|x + y|_v \leq |x|_v + |y|_v$$

pour touts $x, y \in K$ et nous disons que la valuation de ce type est archimedeienne. Si la condition est satisfait avec $C = 1$, alors $|\cdot|_v$ satisfait l'inégalité triangulaire plus forte (ultrametrique):

$$|x + y|_v \leq \max(|x|_v, |y|_v)$$

pour touts $x, y \in K$ en nous disons que cette valuation est non-archimedienne. Ici nous travaillons simplement avec les valuations non-archimediennes, parce que dans le cas de $\mathbb{F}_q[T]$ nous n'avons pas des valuations archimediennes (toutefois il existe une valuation spécifique qui est similaire avec la valuation archimedienne, mais ne peut pas creer quelques problemes pour nous).

Pour chaque v notons par $\mathcal{O}_{(v)}$ l'anneau de fonctions rationnels sur $\mathcal{C}$ qui sont réguliers en v. Dans notre cas ($\mathcal{C} = \mathbb{P}^1$) les valuations fini correspond aux polynômes unitaires irréductible $f \in K = \mathbb{F}_q[T]$. Si cette valuation v correspond à f, alors

$$\mathcal{O}_{(v)} = \{g/h. \text{ t.q. } g, h \in K, \deg(g) < \deg(h)\}.$$

Supposons que le degré de $v = \infty$ est 1. Écrivons $\mathcal{M}_v \subset \mathcal{O}_{(v)}$ pour l'idéal maximal (les éléments de $\mathcal{M}_v$ sont les fonctions qui sont 0 en v) et $\kappa_v = \mathcal{O}_{(v)}/\mathcal{M}_v$ pour le corps résiduel en v. Définissons $\deg(v) = [\kappa_v : k]$, $q_v = q^{\deg(v)}$ pour la norme de v. Maintenant nous devons choisir le modelé minimal entier pour E qui est donnée par l'équation de Weierstrass. Soit $\bar{a}_i \in \kappa_v$ les réductions de coefficients en v. Définissons la courbe réduite E_v par

$$E_v : \ y^2 + \bar{a}_1 xy + \bar{a}_3 y = x^3 + \bar{a}_2 x^2 + \bar{a}_4 x + \bar{a}_6 \tag{2}$$

sur le corps résiduel κ_v. Disons que E_v a

- bonne réduction en v si E_v défini une autre courbe elliptique sur κ_v ($v \nmid \Delta$),
- réduction multiplicative (nodal) en v si E_v a un node en v. Dans ce cas-la il y a deux possibilités différentes. Si les lignes de tangentes dans le node sont rationnelles sur le corps résiduel κ_v, alors disons que ce type de réduction est split multiplicative. Sinon non-split multiplicative.
- réduction additive (cuspidal) en v si E_v a un cusp en v.

Notons que les mots multiplicative et additive sont utilise ici pour marquer que la partie non-singulière de la courbe réduite, qui est défini par

$$E_v^* = E_v/\{\text{un point singulier}\}$$

est isomorphe à $\mathbb{G}_m$ (ou $\mathbb{G}_m[\cdot]$ dans le cas de réduction non-split multiplicative) et $\mathbb{G}_a$ respectivement (ici $\mathbb{G}_m$ est un groupe multiplicative, $\mathbb{G}_m[\cdot]$ est un groupe tordu multiplicative et $\mathbb{G}_a$ est un groupe additive). Les courbes elliptiques, $\mathbb{G}_a$, $\mathbb{G}_m$ et $\mathbb{G}_m[\cdot]$ sur K sont uniquement les courbes algébriques irréductibles sur K, qui ont une structure du groupe, qui est donnée par les fonctions régulières.

La courbe réduite E_v peut-être singulière, mais l'ensemble de points non singulières de $\tilde{E}_v(K_v)$ est toujours un groupe. En plus $E(K)$ a la filtration de groupes abéliennes suivant

$$E_1(K) \subset E_0(K) \subset E(K),$$

où

$$E_0(K) = \{P \in E(K) : P_v \in \tilde{E}_v(K_v)\}$$

et

$$E_1(K) = \{P \in E(K) : P_v = O_v\}$$

avec P_v qui est l'image de $P \in E(K)$ sous un map de réduction $E(K) \to \tilde{E}_v(K_v)$.

Le modelé pour E est donne par E_v avec leur coefficients $\bar{a}_i \in \mathcal{O}_{(v)}$ est appelle entier en v. Le modelé minimal entier en v est un modelé E_v avec le valuation de discriminant Δ de E minimal. L'exponentiel local n_v du conducteur de v est donne par

$$n_v = \begin{cases} 0, & \text{si } E \text{ a bonne réduction en } v, \\ 1, & \text{si } E \text{ a réduction multiplicative en } v, \\ 2 + \delta_v, & \text{si } E \text{ a réduction additive en } v, \end{cases}$$

où δ_v est la ramification sauvage

$$\delta_v = \begin{cases} 0, & \text{si } p > 3, \\ \geqslant 0, & \text{si } p = 2, 3. \end{cases}$$

Alors n_v donne information sur le ramification dans les extensions du corps qui sont générées par les points d'ordre fini dans le sensé de la loi du groupe de la courbe elliptique E.

Le conducteur de E/K est donne par le produit de l'idéel premiers et les exponentielles associées n_v. Le conducteur(global) de E est un diviseur

$$N = \sum_v n_v [v].$$

Le degré du conducteur est

$$\deg N = \sum_v n_v \deg v.$$

N est un diviseur effectif sur $\mathbb{P}^1$ qui est divise seulement par les places v de la réduction mauvaise de E. Le fonction L de E est défini par le produit d'Euler

$$L(E, s) = \prod_{v \nmid N}^{\text{bonne}} \left(1 - \frac{a_v}{q_v^s} + \frac{q_v}{{q_v}^{2s}}\right)^{-1} \times \prod_{v | N}^{\text{mult}} \left(1 - \frac{1}{q_v^s}\right)^{-1} \tag{3}$$

où "bonne" est écrit pour "E a bonne réduction en v", "mult" – pour le cas de réduction split multiplicative ou non split multiplicative en v et, finalement, a_v est un entier défini comme

$$a_v = \begin{cases} q_v + 1 - \#E_v(k_v), & \text{si } E \text{ a bonne réduction en } v, \\ \pm 1, & \text{si } E \text{ a réduction multiplicative en } v, \\ 0, & \text{si } E \text{ a réduction additive en } v. \end{cases}$$

($a_v = 1$ pour réduction split multiplicative et $a_v = -1$ pour réduction non split multiplicative). Par le résultat de Hasse sur la borne de a_v le produit premier de (3) est convergent absolument pour $\operatorname{Re} s > 3/2$ est il y a une continuation méromorphe sur $\mathbb{C}$. Comme d'habitude nous définissons le rang analytique de E/K comme l'ordre d'annulation de la fonction L en $s = 1$

$$\operatorname{rank}_{an}(E) = \operatorname{ord}_{s=1} L(E, s).$$

Rappellos que la courbe elliptique E/K est dit constante si elle peut être définie par l'équation de Weierstrass avec les coefficients appartiennent de k. Elle est dit non-constante si c'est pas constante. Aussi, E/K est dit isotrivielle si ça définie la courbe constante sur quelque extension fini de K, sinon – non-isotrivielle.

Remarque. *Dans le cas non-constante de E Théorème 9.3 de [46] nous donne une borne supérieure du type* $\operatorname{rank}_{an} E \leqslant N$.

La conjecture très célébré de Birch et Swinnerton-Dyer connecte les fonctions L de courbes elliptiques avec le groupe de K-points rationnels sur E/K, en particulier (parmi d'autres relations) elle prévoit que

$$\operatorname{rank}_{an}(E) \stackrel{?}{=} \operatorname{rank}(E).$$

Alors que la conjecture originale reste ouvert, beaucoup plus est connu dans ce contexte du cas des corps de fonctions.

Théorème (Tate [43], Milne [31]). *Soit E est une courbe elliptique sur le corps de fonctions K. Alors*

$$\operatorname{rank} E \leq \operatorname{rank}_{an} E. \tag{4}$$

La technique habituelle pour obtenir des limites supérieures d'un rang analytique est d'utiliser la formule explicite. Nous nous référons ici pour le résultat donné par [5].

Théorème (Brumer [5]). *Soit E est une courbe elliptique sur $\mathbb{F}_q[T]$. Alors le rang analytique est borné par*

$$\operatorname{rank}_{an} E \leq \frac{(b_E - 4)\log q}{2\log b_E} + O\left(\frac{b_E \log^2 q}{\sqrt{q}\log^2 b_E}\right), \tag{5}$$

où b_E est le degré de la fonction L comme un polynomial en q^{-s}.

Pour le cas de $\mathbb{F}_q[T]$ nous avons

$$b_E = n_E - 4,$$

où $n_E = \deg N$ et N est un conducteur de la courbe E/K.
Notons que si E a a réductions additives et m réductions multiplicatives, alors

$$n_E \geq 2a + m.$$

Ce résultat est interessant si et seulement si n_E est assez grand, parce que

$$n_E + 4g_X - 4$$

est une borne trivivale pour le rang. Alors nous avons

$$\operatorname{rank}_{an} E \leq \frac{(\deg N - 8)\log q}{2\log \deg N} + O\left(\frac{\deg N \log^2 q}{\sqrt{q}\log^2 \deg N}\right). \tag{6}$$

La borne simple est

$$\operatorname{rank} E \leq \operatorname{rank}_{an} E \leq b_E = n_E - 4.$$

Si E est constante, alors $\operatorname{rank} E = 0$.

Les propriétés basiques de hauteurs local et canonique.

Maintenant nous parlons un petit peu sur la méthode que nous allons utiliser pour la preuve de Théorème 3. L'idée générale est de développer les bornes pour les hauteurs canonique et local et d'après ça appliquer le résultat d'empilement de sphères.

Dans cette section nous voulons investiguer quelques propriétés de la fonction de hauteur sur la courbe elliptique E sur un corps K. Le fait fondamentale ici est que $|\hat{h} - \frac{1}{2}h_x|$ et $|\hat{h}^E - \frac{1}{3}h_y|$ sont bien bornées sur l'ensemble de tous les points de E. Cela nous permet de donner une borne inférieur pour $\hat{h}^E(P)$ et aussi d'estimer le nombre de points avec la condition $\hat{h}^E < c_2$ sous le restriction que E n'a pas les points tordues P tel que $\hat{h}^E(P) < c_1$. Toutefois, ce chemin nous conduit à un problème que la borne dépendrait de la courbe. Pour éviter cette difficulté, nous allons utiliser des hauteurs locales comme dans [18] et obtenir une borne

$$\lambda_v(P - Q) \geq \min(\lambda_v(P), \lambda_v(Q)),$$

qui marche pas seulement dans le cas de réduction mauvaise avec laquelle nous traiterons séparément. Nous allons subdiviser $E(K_v)$ dans un nombre suffisamment petit de tranches, de sorte que

$$\lambda_v(P-Q) \geq \min(\lambda_v(P), \lambda_v(Q))$$

est toujours vrai sur ces tranches avec P, Q appartiennent de la même tranche (voir Lemme 4 et Lemme 2). L'utilisation que nous allons prouver que les points entiers que nous voulons compter sommes éloignés les uns des autres dans le réseau Mordell -Weil. Rappelons que toute courbe elliptique sur K peut être écrit sous la forme suivante

$$E: \ y^2 = f(x), \tag{7}$$

où $f(x) \in K$ est un polynôme cubique qui est défini par l'équation de Weierstrass. Disons que $d \in K$ est libre de carré si et seulement si il 'y pas de facteur de la forme g^2 avec $g \in K$ et $\deg g \geq 1$. Pour tout $d \in K$ libre de carré définissons le twist quadratique de E comme

$$E_d: \ dy^2 = f(x). \tag{8}$$

Notons que nous nous limitons au cas du d libre de carré, car si d a un facteur carré, puis par un changement de variables, nous pouvons trouver une courbe E_d^* qui est isomorphe à E_d. Écrivons $\hat{h}^E$ pour le hauteur canonique sur une courbe elliptique E, et h_x, h_y pour le hauteur sur E en x et y:

$$\hat{h}^E((x,y)) = \lim_{n\to\infty} \frac{1}{n^2} h_x([n](x,y)), \tag{9}$$

où nous utilisons la notation $[n]P = \underbrace{P + \ldots + P}_{n \text{ fois}}$ et

$$h_x((x,y)) = \begin{cases} 0, & \text{si } P = \mathcal{O}, \\ \log_q H(x), & \text{sinon}, \end{cases}$$

$$h_y((x,y)) = \begin{cases} 0, & \text{si } P = \mathcal{O}, \\ \log_q H(y), & \text{sinon.} \end{cases}$$

Pour tout $x \in K$ définissons la norme de x par $|x| = q^{\deg x}$. Notons que $\hat{h}^E$ est défini pour tous les points de $E(\bar{K})$ et $\hat{h}$ est une forme quadratique définie positive sur $E(\bar{K})$ et bien aussi sur $E(K)$(dans le sens qu'il mappe les éléments non-torsion à des nombres positifs).

Pour $x = x_0/x_1$ avec $x_0, x_1 \in K$ sans un facteur commun autre qu'un polynôme constant K (notons ce fait par $(x_0, x_1)_K = 1$), c'est possible d'écrire $H(x) = \max(|x_0|, |x_1|)$. Soit L est quelque extension algébrique du corps $\mathbb{F}_q[T]$. Définissions $H(y)$ par

$$H(y) = (H_L(y))^{[L:K]^{-1}},$$
$$H_L(y) = \prod_w \max(|y|_w^{n_w}, 1),$$

où $y \in L$, le produit est pris en charge tous les valuations w de L, n_w représente le degré du corps résiduel $L_w/K_w[T]$. Par exemple, si $y = \frac{y_0}{y_1}$ avec $y_0, y_1 \in K$, alors $y \in \mathbb{F}_q(T)$ et pour $L = \mathbb{F}_q(T)$ nous avons

$$H(y) = H_L(y) = \max(|y_0|, |y_1|).$$

Nous énumérons quelques propriétés importantes de la hauteur canonique dans le lemme suivant.

Lemme 1. *Soit $f(x) \in K$ est un polynôme cubique unitaire de discriminant non-zero. Soit aussi d est un polynôme libre de carre $d \in K$ et $P = (x, y)$ est un K-point sur un twist quadratique E_d de E. Soit $P' = (x, d^{1/2}y)$ est un point sur $E_1 = E$ qui est associe à P. Alors*

1. *$\hat{h}^{E_d}(P) = \hat{h}^E(P')$, où les hauteurs canoniques sont définies sur E_d et E respectivement.*

2. *La hauteur h_y $(y \neq 0)$ est bornée sur E, en particulier $h_y(P') \geq \frac{3}{8} \deg d$.*

3. *Nous avons*
$$\hat{h}^{E_d}(P) \geq \frac{1}{8} \deg d + c_f,$$
où c_f est une constante qui dépend seulement de f.

Dans notre preuve nous allons utiliser la propriété suivante.

Corollaire 1. *Soit E est une courbe elliptique sur K. S'il n'y a pas de points non-torsion $P \in E(K)$ de la hauteur canonique $\hat{h}(P) < c_1$, alors il y a au plus*

$$O\left(\left(1 + 2\sqrt{\frac{c_2}{c_1}}\right)^{\operatorname{rank} E}\right)$$

de points dans $E(K)$ de la hauteur canonique $< c_2$.

Notons que le Corrolaire ci-dessus donne une borne uniforme pour $\#E(K)_{tors}$.

Les constantes implicites c_1, c_2 n'ont pas de dépendance de la torsion, mais dépendent de la courbe. Cela nous amène à un problème si nous voulons lié la hauteur canonique en termes de hauteur naïve (à savoir, nous voulons quelque chose de la sorte $h(P) \leq c_3$, où $h(P)$ est la hauteur naïve et c_3 est une constante absolue), parce que dans ce cas la constante dans le grand O changera à $(1 + 2\sqrt{c_3/c_1})^{\operatorname{rank} E}$, où c_1 dépend seulement de la courbe, tandis que c_3 dépend de tous les deux – la courbe et c_2 (disons, $c_2 = c_3 + O_E(1)$). Pour éviter cette difficulté il faut exclure la dépendance masquée par la méthode proposée dans [23].

Rappellons que κ_v est un corps résiduel en v et $d_v = \deg(v) = [\kappa_v : k]$. Soit M_k est un ensemble de places v sur K. Pour chaque place $v \in K$, il y a une fonction de la hauteur naturelle λ_v tel que la hauteur canonique sur E peut être donne dans le sens de λ_v

$$\hat{h}^E(P) = \frac{1}{[K : \mathbb{Q}]} \sum_{v \in M_K} d_v \lambda_v(P).$$

Disons que la courbe elliptique E sur un corps local non-archimédien K a réduction potentiellement bonne si il y a un modelé avec bonne réduction dans quelque extension de K. Au même façon, disons que E a réduction potentiellement multiplicative si il n'y pas de réduction potentiellement bonne. Le lemme suivant est le résultat de [24].

Lemme 2 (Helfgott-Venkatesh [24]). *Soit E est une courbe elliptique sur un corps local non-archimédien K_v avec une réduction potentiellement multiplicative. Alors pour chaque $\varepsilon > 0$ suffisamment petit, il y a un subdivision*

$$E(K_v) = W_{v,0} \cup W_{v,1} \cup \ldots \cup W_{v,d_v}$$

avec $d_v \ll |\log \varepsilon|$, tel que pour tous les deux points distincts $P, Q \in W_{v,0}$ nous avons

$$\lambda_v(P - Q) \geq \min(\lambda_v(P), \lambda_v(Q)), \quad \lambda_v(P_1), \lambda_v(P_2) \geq 0,$$

et pour tous les deux points distincts $P, Q \in W_{v,j}$, où $1 \leq j \leq d_v$ nous avons

$$\begin{aligned} \lambda_v(P - Q) &\geq (1 - \varepsilon) \max(\lambda_v(P), \lambda_v(Q)), \\ \lambda_v(P - Q) &\geq (1 - 2\varepsilon) \max(\lambda_v(P), \lambda_v(Q)), \end{aligned}$$

où la constante implicite est absolue.

Nous allons exploiter Lemme 5 pour donner une limite supérieure sur le nombre de S-points intégrales. Pour obtenir une bonne constante C, qui a comparu devant Théorème 3, nous allons appliquer l'empilement de sphères. Nous divisons d'abord l'ensemble des points entiers sur E en 'bonnes tranches' et appliquer des limites d'empilement de sphères à chaque partie séparément. Ici, nous utilisons le résultat remarquable de Kabatiansky et Levenstein [27].

Lemme 3 (Kabatiansky-Levenstein [27]). *Soit $A(n, \theta)$ est un nombre maximal de points qui peut être disposé sur la sphère unitaire de $\mathbb{R}^n$ de telle sorte que l'angle entre P_1, O et P_2 pour tous les deux points P_1, P_2 n'est pas plus petit que θ. Alors pour $0 < \theta < \frac{\pi}{2}$ nous avons*

$$\frac{1}{n} \log_2 A(n, \theta) \leq \frac{1 + \sin\theta}{2\sin\theta} \log_2 \frac{1 + \sin\theta}{2\sin\theta} - \frac{1 - \sin\theta}{2\sin\theta} \log_2 \frac{1 - \sin\theta}{2\sin\theta} + o(1),$$

où la convergence est uniforme et explicite pour θ apartient de chaque sous-intervalle fermé de $\left(0, \frac{\pi}{2}\right)$. En particulier, pour $\theta = \frac{\pi}{3}$, nous avons

$$\frac{1}{n} \log_2 A(m, \theta) \leq 0.40141 \ldots.$$

Brièvement, nous subdivisons les points S-entiers sur E, que nous notons par $E(K, S)$ en points mod I pour I étant un idéal adapté en O_K. Puis Lemme 5 dit que, après quelques manipulations sur cette partition les points, qui se trouvent dans la même classe ont tendance à être éloignés les uns des autres dans le réseau Mordell-Weil. A cette point nous appliquons Lemme 3 à chaque partie séparément. Ces bornes d'"empilement de sphères vont nous amener à terme $e^{\beta(t) \operatorname{rank} E}$ dans chaque partie. Sommation sur toutes les classes donne lieu à un autre terme $e^{[K:L]h_0}$. Il suffit simplement de prendre soin d'obtenir les bonnes conditions pour appliquer Lemme 5.

1.2 La partie analytique

Le théorème de Bombieri-Vinogradov [12, Chapitre 28] (pour la discussion importante sur le sujet voir aussi [36]) est un résultat très connu dans la théorie analytique de nombres. Ce théorème concerne le terme d'erreur dans le théorème de Dirichlet sur les progressions arithmétiques, pris en moyenne sur les modules q variant jusqu'à Q. Parfois ce s'appelle l'hypothèse de Riemann generalisee (GRH), parce que sans faire la moyenne, c'est environ de la force de GRH.

Nous commençons par quelques définitions nécessaires. Pour un entier a et $q \geq 1$ soient

$$\psi(x;q,a) = \sum_{\substack{n\leq x \\ n\equiv a \bmod q}} \Lambda(n), \quad \psi(x;1,a) = \psi(x),$$

où $\Lambda(n)$ est une fonction de von Mangoldt, qui est défini par

$$\Lambda(n) = \begin{cases} \log p, & \text{si } n = p^k \text{ pour un nombre premier } p \text{ et un entier } k \geq 1, \\ 0, & \text{sinon.} \end{cases}$$

Notons par $\varphi(q)$ l'indicatrice d'Euler, qui est le nombre de termes pour le module q

$$\varphi(n) = n \prod_{p|n} \left(1 - \frac{1}{p}\right).$$

Le théorème de Dirichlet sur les progressions arithmétiques dit que

$$\psi(x;q,a) \sim \frac{x}{\varphi(q)},$$

comme $x \to \infty$. Le théorème de Bombieri-Vinogradov est le résultat suivant.

Théorème (Bombieri-Vinogradov)**.** *Soit $A > 0$ et $Q \leq \frac{x^{\frac{1}{2}}}{(\log x)^B}$, où $B = B(A)$. Alors*

$$\sum_{q\leq Q} \max_{2\leq y\leq x} \max_{\substack{a \\ (a,q)=1}} \left|\psi(y,q,a) - \frac{y}{\varphi(q)}\right| \ll_A \frac{x}{(\log x)^A}.$$

La constante implicite n'est pas effective, parce que nous devons travailler avec les caractères associés à telles q, qui ont les petits facteurs premiers (où nous devons utiliser le théorème de Siegel-Walfisz).

Pour un q fixe, en supposant GRH, nous obtenons

$$\psi(x;q,a) - \frac{\psi(x)}{\varphi(q)} = O\left(\sqrt{x}(\log x)^2\right).$$

Alors le théorème de Bombieri-Vinogradov nous donne un resultat en moyenne de la même magnitude que GRH.

Notons que pour Q assez grande la borne dans le théorème de Bombieri- Vinogradov est triviale. Pour $1 \leq y \leq x$ et $q \leq x^{\frac{1}{2}}$ en appliquant l'inégalité de Brun-Titchmarch nous avons

$$\psi(y;q,a) \leq \psi(x;q,a) = \sum_{\substack{n=p^k\leq x \\ n\equiv a (\mathrm{mod}\, q)}} \log n \ll \frac{1}{\varphi(q)} \frac{\log x}{\log \frac{x}{q}} \ll \frac{x}{\varphi(q)}.$$

Cela nous amène à la borne triviale dans le théorème de Bombieri-Vinogradov

$$\sum_{q\le Q}\max_{2\le y\le x}\max_{\substack{a\\(a,q)=1}}\left|\psi(y;q,a)-\frac{y}{\varphi(q)}\right|\ll\sum_{q\le Q}\frac{x}{\varphi(q)}\ll x(\log Q)\ll x(\log x).$$

La constante $B=B(A)$ dans le théorème de Bombieri-Vinogradov initielle était $B=3A+23$, mais la valeur de B a été améliorée. Par exemple, dans [13] Dress, Iwaniec et Tenenbaum ont prouve que $B=A+\frac{5}{2}$ est possible. Bien que [13] n'indique pas le résultat en entier - en se concentrant sur l'estimation de la somme cruciale - une forme complète peut être trouvée dans [1] avec une version entièrement explicite. Au même temps, des versions effiectives – dans lesquelles l'effet d'un caractère exceptionnel est évité d'une manière ou d'une autre - ont été connus depuis que [30] et [45].

Dans cette partie nous améliorons le résultat de [1] et [13] en obtient la puissance du facteur logarithmique $(\log x)^{\frac{7}{2}}$ (Théorème 4) et $(\log x)^2$ (Théorème 5) par contre $(\log x)^{\frac{9}{2}}$ dans [1] et $(\log x)^{\frac{5}{2}}$ dans [13].

Nous formulons notre premier résultat.

Théorème 4 (Le théorème de Bombieri-Vinogradov avec les constantes explicites)**.** *Soit $x\ge 4$, $1\le Q_1\le Q\le x^{\frac{1}{2}}$. Soit aussi $l(q)$ est le plus petit diviseur de q. Notos par $F(x,Q,Q_1)$ le fonction, qui est defini par*

$$F(x,Q,Q_1)=\frac{14x}{Q_1}+4x^{\frac{1}{2}}Q+15x^{\frac{2}{3}}Q^{\frac{1}{2}}+4x^{\frac{5}{6}}\log\frac{Q}{Q_1}.$$

Alors

$$\sum_{\substack{q\le Q\\ l(q)>Q_1}}\max_{2\le y\le x}\max_{\substack{a\\(a,q)=1}}\left|\psi(y;q,a)-\frac{\psi(y)}{\varphi(q)}\right|<c_1F(x,Q,Q_1)(\log x)^{\frac{7}{2}},$$

où

$$\begin{aligned}
c_1&=\frac{5}{4}E_0c_0+1=42.140461\ldots,\\
E_0&=\prod_p\left(1+\frac{1}{p(p-1)}\right)=1.943596\ldots,\\
c_0&=(2A_0)^{\frac{1}{2}}\frac{2^5}{3^{\frac{3}{2}}\pi(\log 2)^2}\left(2+\frac{\log(\log 2)}{\log\frac{4}{3}}\right)=16.93375\ldots,\\
A_0&=\max_{x>0}\left(\frac{\psi(x)}{x}\right)=\frac{\psi(113)}{113}=1.03883\ldots.
\end{aligned}$$

Corollaire. *Soit $A>0$ et $Q\le\frac{x^{1/2}}{(\log x)^A}$. Alors*

$$\sum_{q\le Q}\max_{2\le y\le x}\max_{\substack{a\\(a,q)=1}}\left|\psi(y,q,a)-\frac{y}{\varphi(q)}\right|\ll_A\frac{x}{(\log x)^{A-\frac{7}{2}}}.$$

La constante impliquée ne pas effective.

En utilisant Lemme 4 et Lemme 6 avec le technique de [47] nous nous améliorons et formulons le résultat principal de cette section.

Théorème 5 (Le théorème de Bombieri-Vinogradov, effective)**.** *Soit $x \geq 4$, $1 \leq Q_1 \leq Q \leq x^{\frac{1}{2}}$. Soit aussi $l(q)$ est le plus petit diviseur de q. Pour chaque $\varepsilon < \frac{1}{28}$ positif nous avons*

$$\sum_{\substack{q \leq Q \\ l(q) > Q_1}} \max_{2 \leq y \leq x} \max_{\substack{a \\ (a,q)=1}} \left| \psi(y;q,a) - \frac{\psi(y)}{\varphi(q)} \right| \ll x^{\frac{1}{2}} Q (\log x)^2 + \frac{x}{Q_1} (\log x)^3 + x^{\frac{13}{14}+\varepsilon} (\log x)^4.$$

La constante impliquée peut être faite explicite en utilisant les résultats de [22].

Corollaire 2. *Soit $A > 0$ et $Q \leq \frac{x^{1/2}}{(\log x)^A}$. Alors Proposition 2 nous donne une borne*

$$\sum_{q \leq Q} \max_{2 \leq y \leq x} \max_{\substack{a \\ (a,q)=1}} \left| \psi(y,q,a) - \frac{y}{\varphi(q)} \right| \ll_A \frac{x}{(\log x)^{A-2}}.$$

La constante impliquée dans Corollaire 2 ne pas effective.

L'amélioration ici consiste à avoir un facteur de $(\log x)^2$, plutôt que $(\log x)^{\frac{5}{2}}$ ou $(\log x)^3$ dans l'inégalité de Vaughan. Afin de prouver Proposition 1, nous utilisons la version pondérée de l'identité de Vaughan (voir Lemme 4) et une estimation de façon Barban-Vehov [2] et Graham [17]. Alors que Graham utilise le théorème de Siegel-Walfisz, il existe une version efficace (et explicite) de ce résultat dans [22]. Proposition 1 nous fournit une preuve du théorème Bombieri-Vinogradov. En plus la preuve de Proposition 2 utilise le théorème de Siegel-Walfisz, qui indique que

$$\psi(x,\chi) - \delta(\chi) x \ll_A x e^{-c\sqrt{\log x}}$$

uniformément pour $q \leq (\log x)^A$. Ici $A > 0$ est un nombre réel fixe, c est une constante positive absolue, et $\delta(\chi) = 1$ si χ est le principal et est nul autrement.

La constante implicite n'est pas effective, parce que nous utilisons le théorème de Siegel-Walfisz.

Remarque 1. *Definions*

$$\pi(x) = \sum_{p \leq x} 1 \qquad \textit{and} \qquad \pi(x;q,a) = \sum_{\substack{p \leq x \\ p \equiv a \bmod q}} 1.$$

Alors Théorème 5 sous les mêmes hypothèses peut être aussi formulé pour $\pi(x)$, $\pi(x;q,a)$:

$$\sum_{\substack{q \leq Q \\ l(q) > Q_1}} \max_{2 \leq y \leq x} \max_{\substack{a \\ (a,q)=1}} \left| \pi(y;q,a) - \frac{\pi(y)}{\varphi(q)} \right| \ll x^{\frac{1}{2}} Q (\log x)^2 + \frac{x}{Q_1} (\log x)^3 + x^{\frac{13}{14}+\varepsilon} (\log x)^4.$$

La preuve de Remarque 1 est exactement le même que dans [1], il suffit de changer la puissance de $(\log x)$.

Les outils qui nous permettions à procéder sont l'inégalité de grand crible, l'identité de Vaughan et le lemme pour l'estimation la contribution de la fonction Möbius

$$\mu(n) = \begin{cases} (-1)^{\omega(n)}, & \text{si } p^2 \nmid n \text{ pour chaque nombre premier } p, \\ 0, & \text{sinon,} \end{cases}$$

où $\omega(n)$ est le nombre de facteurs premiers de n sans les multiplicités.

La méthode du grand crible.

Le grand crible commence par un travail de Linnik en 1941. Il était développé par Rényi à partir de 1947, mais resta très complexe jusqu'à 1975 où de nombreux travaux simplifierent beaucoup la formulation. En particulier, en 1966, Davenport et Halberstam trouvent une simple inégalité analytique très générale, qui s'appelle maintenant le principe analytique du grand crible (voir, par example, [33, p.561]).

Définissons un caractère de Dirichlet χ comme une fonction de l'ensemble $\mathbb{N}^*$ des entiers strictement positifs dans $\mathbb{C}$, totalement multiplicative et périodique. Si n est la période et d un entier strictement positif, $\chi(d)$ est de module 1 si d est premier avec n et nul sinon.

Nous donnons ici le principe analytique du grand crible dans la forme de Gallacher [16].

Lemme (Le principe analytique du grand crible)**.** *Soit $Q > 0$, a_m est une séquence arbitraire de nombres complexes. Alors*

$$\sum_{q\le Q}\frac{q}{\varphi(q)}\sideset{}{^*}\sum_{\chi(q)}\left|\sum_{m=m_0+1}^{m_0+M}a_m\chi(m)\right|^2 \le (M+Q^2)\sum_{m=m_0+1}^{m_0+M}|a_m|^2. \tag{10}$$

Ici $\sum^*$ désigne une somme sur tous les caractères primitifs $\chi(\mathrm{mod}\, q)$. D'après le résultat ci-dessus c'est facile de prouver (voir [1, Lemme 6.1] et aussi [9, Chapter 8.3])

Le lemme suivant est la modification nécessaire du principe du grand crible.

Lemme. *Soit a_m, b_n sont les séquences arbitraires de nombres complexes. Alors*

$$\sum_{q\le Q}\frac{q}{\varphi(q)}\sideset{}{^*}\sum_{\chi(q)}\max_y\left|\sum_{m=m_0}^{M}\sum_{\substack{n=n_0\\ mn\le y}}^{N}a_mb_n\chi(mn)\right| \le$$

$$c_3(M'+Q^2)^{\frac12}(N'+Q^2)^{\frac12}\left(\sum_{m=m_0}^{M}|a_m|^2\right)^{\frac12}\left(\sum_{n=n_0}^{N}|b_n|^2\right)^{\frac12}L(M,N), \tag{11}$$

où $c_3 = 2.64...$, et

$$L(M,N) = \log(2MN),$$
$$M' = M - m_0 + 1,$$
$$N' = N - n_0 + 1.$$

(M', N' sont les nombres de termes dans les sommes sur m et n respectivement.)

C'est clair que ce n'est pas possible d'obtenir la version du lemme ci-dessus sans un facteur logarithmique $L(M,N)$, parce que nous avons la condition multiplicative $mn \le y$.

L'inégalité de Vaughan.

L'identité de Vaughan est une identité trouvée par R. Vaughan en 1977 (voir [48]), qui peut être utilisé pour simplifier le travail de Vinogradov sur les sommes trigonométriques [49, Chapitre IX]. Il peut être utilisé pour estimer les sommes du type $\sum f(n)\Lambda(n)$, où f est une fonction d'entiers positifs n, dont les valeurs dans les applications sont souvent des racines d'unité.

Lemme (L'identité de Vaughan)**.** *Soit U, V sont deux paramètres. Alors*

$$\Lambda(n) = \lambda_0(n) + \lambda_1(n) + \lambda_2(n) + \lambda_3(n),$$

où

$$\begin{aligned}
\lambda_0(n) &= \begin{cases} \Lambda(n), & si\ n \le U, \\ 0, & si\ n > U, \end{cases} \\
\lambda_1(n) &= \sum_{\substack{hd=n \\ d \le V}} \mu(d) \log h, \\
\lambda_2(n) &= -\sum_{\substack{mdr=n \\ m \le U, d \le V}} \Lambda(m)\mu(d), \\
\lambda_3(n) &= -\sum_{\substack{mk=n \\ m > U, k > V}} \Lambda(m) \sum_{\substack{d|k \\ d \le V}} \mu(d).
\end{aligned}$$

Le lemme ci-dessus est suffisant pour prouver le Proposition 1 et donc le Théorème 4. Nous prouvons aussi la version pondérée de cette inégalité, dont nous avons besoin dans la preuve du Théorème 5.

Lemme 4 (L'identité de Vaughan pondérée)**.** *Soit $U, V \ge 1$. Definissons une fonction $\eta(t) : \mathbb{Z}^+ \to \mathbb{R}$, qui satisfait $\eta(t) = 1$ pour $t \le V$. Nous avons*

$$\Lambda(n) = \lambda_0(n) + \lambda_1(n) + \lambda_2(n) + \lambda_3(n),$$

où

$$\begin{aligned}
\lambda_0(n) &= \begin{cases} \Lambda(n), & si\ n \le U, \\ 0, & si\ n > U, \end{cases} \\
\lambda_1(n) &= \sum_{d|n} \mu(d)\eta(d) \log \frac{n}{d}, \\
\lambda_2(n) &= -\sum_{c \le U} \sum_{dc|n} \mu(d)\Lambda(c)\eta(d), \\
\lambda_3(n) &= \sum_{c > U} \sum_{dc|n} \mu(d)\Lambda(c)(1 - \eta(d)).
\end{aligned}$$

Dans notre preuve nous utilisons cette inégalité et obtenons une version explicite qu'on appelle l'inégalité de Vaughan explicite (voir Proposition 2). C'est une partie importante pour la preuve de théorème de Bombieri Vinogradov. Nous pouvons utiliser les autres inégalités, par exemple, l'inégalité de Daboussi ([10] et [11]) ou l'inégalité de Heath-Brown [20], mais ça va prolonger les calculs sans succès facile.

La fonction de Möbius.

Soit V est un parametre. Definissons

$$b_k = \sum_{\substack{d \le V \\ d|k}} \mu(d),$$

où $\mu(d)$ est le fonction de Möbius. Le somme de ce type était étudié par les auteurs dans [13], où ils ont prouvé que

$$\sum_{d_1,d_2\leq Y}\frac{\mu(d_1)\mu(d_2)}{\gcd(d_1,d_2)}$$

tend vers une constante positive quand $Y\to\infty$. Il est également suggéré sans prouver que L peut être egal à environ 0.440729. Nous utilisons la version explicite de la formule ci-dessus, qui était donné par Helfgott, voir [22].

Lemme 5. *Pour V assez grand nous avons*

$$\left|\frac{1}{Y}\sum_{k\leq Y}|b_k|^2-L\right|\leq\frac{V^2}{Y}+C,$$

où $C=0.000023$, $L=0.440729$.

Pour la démonstration du Théorème 5, nous avons besoin de l'estimation du travail de Graham [17].

Lemme 6 (Graham [17])**.** *Soit $1\leq N_1\leq N_2\leq N$ et definissons*

$$f_i(d)=\begin{cases}\mu(d)\log\frac{N_i}{d}, & d\leq N_i,\\ 0, & d>N.\end{cases}$$

Alors nous avons

$$\sum_{n=1}^{N}\left(\sum_{d_1|n}f_1(d_1)\right)\left(\sum_{d_2|n}f_2(d_2)\right)=N\log N_1+O(N).$$

Pour la preuve, voir Graham [17]. À partir du lemme ci-dessus nous pouvons en déduire

Corollaire 3. *Definissons une foction $\eta(t)$, qui est égale à 1 pour $t\leq V$, à 0 pour $t>V_0$ et*

$$\eta(t)=\frac{\log\frac{V_0}{t}}{\log\frac{V_0}{V}},\ \ V<t\leq V_0.$$

Alors

$$\sum_{k\leq Y}\left|\sum_{d|k}\mu(d)\eta(d)\right|^2\ll\frac{Y}{\log\frac{V_0}{V}}.$$

La constante ici peut-être faire explicite grâce à [22].

2 Bombieri-Pila theorem over function fields

2.1 Introduction

In [4] E. Bombieri and J. Pila proved that if Γ is a subset of an irreducible algebraic curve of degree d inside a square of side N, then the number of lattice points on Γ is bounded by $c(d,\varepsilon)N^{\frac{1}{d}+\varepsilon}$ for any $\varepsilon > 0$, where the constant $c(d,\varepsilon)$ does not depend on Γ. There are many analogues of this remarkable result. For example, one can be interested in finding a bound for a number of solutions of $f(x,y) = 0 (\mathrm{mod}\, p)$ with $x \in I$, $y \in J$, where I and J are short intervals in $\mathbb{Z}/p\mathbb{Z}$ (see [6] and [8]). Such results are p-analogues of the Bombieri-Pila bound. (Here we should assume that the lengths of I and J are much shorter than p, so that the Weil bound and other standard methods cannot be applied).

One can go further and look for a function field analogue. Here we work in a finite field $\mathbb{F}_{q^n}$ modelled as $\mathbb{F}_q[T]/f(T)$ where f is a fixed irreducible polynomial of degree n and T is a formal variable. Then one can define an "interval" as a set of polynomials of the form $X + Y = X(T) + Y(T)$, where $X \in \mathbb{F}_q[T]$ is a fixed polynomial and $Y(T) \in \mathbb{F}_q[T]$ runs through all polynomials of degree bounded by a given natural number. This point of view was used by J. Cilleruelo and I. Shparlinski in [7] for obtaining some bounds on the number of solutions of polynomial congruences modulo a prime with variables in short intervals. The same authors also formulated [7, Problem 9], that is solved here.

In what follows we work with a planar irreducible algebraic curve $\mathcal{C}$ of degree d over $K = \mathbb{F}_q[T]$. The first principal result of this section is the following.

Theorem 1. *Let $\mathcal{C}$ be an irreducible algebraic curve of degree d over $\mathbb{F}_q[T]$, q is a prime power. Define S as the set of points on $\mathcal{C}$ inside I^2, where I is a set of polynomials $X \in \mathbb{F}_q[T]$ with $\deg X \leqslant n$ and $|I| = q^{n+1}$. Then*

$$|S| \ll_{d,\varepsilon} |I|^{\frac{1}{d}+\varepsilon}.$$

One can pose a question: why can we not just follow the Bombieri-Pila approach in order to get Theorem 1? Unfortunately, in this case we will cross some difficulties in getting Lemma 2 of [4], since we do not have the necessary analogue of the mean value theorem in function fields, see [42, Lemma 1]. There seem to be at least two plausible ways to avoid this difficulty. The first one consists in getting a function field variant of Theorem 14 in Heath-Brown's article [21]. The second one is to adapt the method of Helfgott-Venkatesh [25]. Since the first one appears to require some long and technical computations we will follow here the second approach.

We will need analogues of Propositions 3.1 and 3.2 of [25]. Combining and developing the original ideas of [4] together with an adaptation of some results of [25] will lead us to our main result.

After that we will use Theorem 1 to get some applications, such as bounds on the number of isomorphism classes which are represented by elliptic curves $E_{a,b}$ parametrized by coefficients $a, b \in \mathbb{F}_q[T]$ lying in a small box, say, I^2. Using this result one can estimate the number of elliptic curves lying in a given isomorphism class with coefficients lying in a small box. To proceed we will work with ideas proposed in [8].

2.2 Auxiliary statements

Let q be a prime power, X and Y be variables with values in $\mathbb{F}_q[T]$, i.e. their values are of the form

$$X = X(T) = a_0 + a_1 T + \ldots + a_n T^n,$$
$$Y = Y(T) = b_0 + b_1 T + \ldots + b_m T^m,$$

where T is a place holder, $a_i, b_j \in \mathbb{F}_q$, $i = 0, \ldots, \deg X = n$, $j = 0, \ldots, \deg Y = m$. For $X \in \mathbb{F}_q[T]$ we denote by $|X|$ its norm: $|X| = q^{\deg X}$.

In what follows $\mathcal{C}$ is an irreducible algebraic curve of degree d over $\mathbb{F}_q[T]$, which is described by an equation $F(X, Y) = 0$, $F(X, Y) \in (\mathbb{F}_q[T])[X, Y]$. Write S for the set of points on $\mathcal{C}$ inside I^2, where I is an interval in $\mathbb{F}_q[T]$ (we fix $X = 0$ in the definition of I).

For any $F(X, Y) \in (\mathbb{F}_q[T])[X, Y]$ we write $\deg_X F$ and $\deg_T F$ to denote the degree of a polynomial F with respect to X and T respectively. We also use the standard notation $\deg F(X, Y)$ for the degree of $F(X, Y)$ as a polynomial in X and Y.

Let $\mathcal{W}$ be a set consisting of finitely many linearly independent monomials $F \in (\mathbb{F}_q[T])[X, Y]$ including the one – $\mathbf{1}$. Write $d_{\mathcal{W}}$ for the sum of degrees of all elements that belong to $\mathcal{W}$ and define $\omega = |\mathcal{W}|$. Assume that the elements of $\mathcal{W}$ separate points, meaning that for any two distinct elements $(X_1, Y_1), (X_2, Y_2)$ of $(\mathbb{F}_q[T])^2$ there is an $F \in \mathcal{W}$ such that $F(X_1, Y_1) \neq F(X_2, Y_2)$. We define a $\mathcal{W}$-curve to be an affine algebraic curve described by an equation $G(X, Y) = 0$, where all the monomials of G belong to $\mathcal{W}$.
During the proof of Theorem 1 we will use the following choice of $\mathcal{W}$:

Example 1. *Define $\mathcal{W} = \mathcal{W}_{d,M}$ as*

$$\mathcal{W} = \{X^i Y^j |\ i \leqslant d, j \leqslant M\},$$

where d and M are given numbers.
Then

$$\omega = |\mathcal{W}| = (d+1)(M+1),$$
$$d_{\mathcal{W}} = \sum_{F \in \mathcal{W}} \deg F = (d+1)(M+1)\frac{d+M}{2}.$$

The $\mathcal{W}$-curves are plane curves of degree less or equal than d and M in X and Y respectively.

This choice is taken straight from the work of Bombieri and Pila [4].

Lemma 1. *Let $\mathcal{C}$ be an irreducible algebraic curve of degree d over $\mathbb{F}_q[T]$ and let S be the set of points on $\mathcal{C}$ inside I^2. Suppose that the number of residues $\{(X, Y) \bmod f,\ X, Y \in S\}$ is at most $\alpha|f|$ for some fixed $\alpha > 0$ and for every irreducible polynomial $f \in \mathbb{F}_q[T]$ with $|f| > c$. Assume that $\mathcal{W}$ is chosen in a way that any $\mathcal{W}$-curve contains at most $\delta|S|$ elements of S, where δ is small enough. Then as $|I| \to \infty$ the following holds*

$$|S| \ll_{\mathcal{W},\delta,c} |I|^{\frac{2\alpha d_{\mathcal{W}}}{\omega(\omega-1)} + O_{\omega,\alpha}(\delta)}.$$

Proof. We are going to prove it in the spirit of [25, Proposition 3.1]. Write $P = (X, Y)$ for a point in $(\mathbb{F}_q[T])^2$ with coordinates $X, Y \in \mathbb{F}_q[T]$. Fixing an arbitrary ordering $F_1, F_2, \dots, F_\omega$ for the elements of $\mathcal{W}$, we define a function

$$W : ((\mathbb{F}_q[T])^2)^\omega \to \mathbb{F}_q[T] \quad \text{by} \quad W(P_1, \dots, P_\omega) = \det(F_i(P_j))_{1 \leqslant i,j \leqslant \omega}.$$

Fix any irreducible polynomial f with $\deg f | \operatorname{ord}_q N$, where N is to be set at the end (we keep in mind that it means that $|f| \leq q^{\operatorname{ord}_q N} \leq N$). Notice that if the number of distinct points among $P_i (\mathrm{mod}\, f)$ is less than or equal to some number k, then

$$\operatorname{ord}_f W(P_1, \dots, P_\omega) \geqslant \omega - k.$$

Let $\mathbb{P}$ denote an ω-tuple of points in S

$$\mathbb{P} = (P_1, \dots, P_\omega), \quad P_i = (X_i, Y_i) \in S.$$

We say that $\mathbb{P}$ is admissible if $W(\mathbb{P}) = W(P_1, \dots, P_\omega) \neq \mathbf{0}$ (where $\mathbf{0}$ stands for the zero polynomial in $\mathbb{F}_q[T]$). Recall that $|X| = q^{\deg X}$ for $X \in \mathbb{F}_q[T]$ and define

$$\Delta = \prod_{\mathbb{P}}^{*} W(\mathbb{P}),$$

where $*$ means that we take the operation over all admissible $\mathbb{P}$.
By the definition of $d_{\mathcal{W}}$ we have

$$|W(\mathbb{P})| \ll_{\mathcal{W}} |I|^{d_{\mathcal{W}}}$$

for every $\mathbb{P} \in S^\omega$. Taking $\log_q |\Delta|$ and applying the expression above gives

$$\frac{\log_q |\Delta|}{|S|^\omega} = \frac{\sum_{\mathbb{P}}^{*} \log_q |W(\mathbb{P})|}{|S|^\omega} \leqslant d_{\mathcal{W}} \log_q |I| + O_{\mathcal{W}}(1). \tag{12}$$

For every point $P \in (\mathbb{F}_q[T])^2$ let ρ_P be the fraction of points in S that reduce to $P \bmod f$. For each $\mathbb{P}$ let $\kappa(\mathbb{P}) \in \{0, 1, \dots, \omega - 1\}$ be defined in a way that $\omega - \kappa(\mathbb{P})$ is the number of distinct points among the points $P_i \bmod f$. Then one can state

$$\operatorname{ord}_f W(\mathbb{P}) \geqslant \omega - (\omega - k(\mathbb{P})) = k(\mathbb{P}),$$

that brings us to

$$\sum_{\mathbb{P}}^{*} \operatorname{ord}_f W(\mathbb{P}) \geqslant \sum_{\mathbb{P}}^{*} \kappa(\mathbb{P}) = \sum_{\mathbb{P}} \kappa(\mathbb{P}) - \sum_{\mathbb{P}}^{na} \kappa(\mathbb{P}), \tag{13}$$

where the first sum on the right hand side is taken over all $\mathbb{P}$ and the second one is the sum over all non-admissible ω-tuples $\mathbb{P}$.

We are going to proceed in two steps. First, we calculate the sum over all $\mathbb{P} \in S^\omega$ by probabilistic methods. Here we see $P_1, \dots, P_\omega$ as ω independent random variables with values in $(\mathbb{F}_q[T])^2$ and use

$$Y_P = \begin{cases} 1, & \text{if at least one of } P_i \neq P \text{ in } S \text{ is equal to } P (\mathrm{mod}\, f); \\ 0, & \text{otherwise.} \end{cases}$$

In the non-admissible case of $\mathbb{P}$ we have either at least two points $P_i = P_j$ among the entries of $\mathbb{P}$ or at least two points $P_i \equiv P_j (\mathrm{mod}\, f)$, $P_i, P_j \in \mathbb{P}$, $i \neq j$. The number of pairs P_i, P_j that satisfy the first possibility can be easily bounded by $O(|S|^{\omega-1})$ and for the latter case we permute the entries of our matrix in order to have

$$\det(F_i(P_j))_{1 \leqslant i,j \leqslant l} \neq 0$$

of a maximal possible size l and then apply the fact that any $\mathcal{W}$-curve contains at most $\delta|S|$ number of elements of S.

Let us start with the sum over all $\mathbb{P} \in S^\omega$. Consider $\mathbb{P}$ as a random variable with uniform distribution. Then the expected value of the number of distinct points among the $P_i \bmod f$ is equal to

$$\frac{\sum_{\mathbb{P}}(\omega - \kappa(\mathbb{P}))}{|S|^\omega} = \mathbb{E}\left(\sum_P Y_P\right).$$

Further,

$$\begin{aligned}
\mathbb{E}\left(\sum_P Y_P\right) &= \sum_P \mathbb{E}(Y_P) = \sum_P \mathrm{Prob}(\exists P_i \neq P | P_i \equiv P \bmod f) \\
&= \sum_P (1 - \mathrm{Prob}(\nexists P_i \neq P | P_i \equiv P \bmod f)) \\
&= \sum_P (1 - \mathrm{Prob}(\forall P_i \neq P | P_i \not\equiv P \bmod f)) \\
&= \sum_P \left(1 - \prod_i \mathrm{Prob}(P_i \not\equiv P \bmod f, P_i \neq P)\right) \\
&= \sum_P \left(1 - \prod_i (1 - \rho_P)\right) \\
&= \sum_P \left(1 - (1 - \rho_P)^\omega\right).
\end{aligned}$$

We then have

$$\frac{\sum_{\mathbb{P}}(\omega - \kappa(\mathbb{P}))}{|S|^\omega} = \sum_P \left(1 - (1 - \rho_P)^\omega\right).$$

Next

$$\frac{\sum_{\mathbb{P}} \kappa(\mathbb{P})}{|S|^\omega} = \frac{\sum_{\mathbb{P}} \omega}{|S|^\omega} - \sum_P (1 - (1 - \rho_P)^\omega) = \sum_P ((1 - \rho_P)^\omega + \omega\rho_P - 1). \tag{14}$$

Now let us bound the sum over all non-admissible $\mathbb{P}$. Consider the set of such $\mathbb{P}$ with $\kappa(\mathbb{P}) > 0$. Then one of the followings is true:

1. There exist $i \neq j$, such that $P_i = P_j$;
2. There exist $i \neq j$, such that $P_i \equiv P_j (\mathrm{mod}\, f)$, but $P_i \neq P_j$.

The total number of non-admissible $\mathbb{P}$, such that the first condition above holds is equal to $O(|S|^{\omega-1})$. Let us estimate this number for the second case. Permute the

entries in such a way that $i = 1$, $j = 2$ and $F_1 = \mathbf{1}$, $F_2(P_i) \neq F_2(P_j)$ (this is possible since we have assumed that the elements of $\mathcal{W}$ separate points and $\mathcal{W}$ contains $\mathbf{1}$). Then for $l = 2$

$$\det(F_i(P_j))_{1 \leqslant i,j \leqslant l} \neq 0.$$

Choose the maximal l, such that the above statement still holds. Then P_{l+1} lies on a $\mathcal{W}$ curve determined by $P_1, P_2, \ldots, P_l$. As we demanded, the number of possible values for P_{l+1} is bounded above by $\delta|S|$. Then the number of non-admissible $\mathbb{P}$, such that the second case takes place is equal to

$$O_\omega(|S|^{\omega-2}\delta r),$$

where r is the number of pairs $(Q_1, Q_2) \in S^2$ that reduce to the same point modulo f. By its definition r can be expressed as

$$r = |S|^2 \sum_P \rho_P^2.$$

Summing two results we see that there are at most

$$O_\omega\left(|S|^{\omega-1} + |S|^\omega \delta \sum_P \rho_P^2\right) = |S|^\omega O_\omega\left(|S|^{-1} + \delta \sum_P \rho_P^2\right) \tag{15}$$

non-admissible $\mathbb{P}$ with $\kappa(\mathbb{P}) > 0$.

We divide (13) by $|S|^\omega$ and insert (14) and (15) divided by $|S|^\omega$ into (13) to see that

$$\sum_{\mathbb{P}}^{*} \frac{\operatorname{ord}_f W(\mathbb{P})}{|S|^\omega} \geq \sum_P ((1-\rho_P)^\omega + \omega\rho_P - 1) + O_\omega\left(|S|^{-1} + \delta \sum_P \rho_P^2\right). \tag{16}$$

Now we have to give an upper bound for the first tern on the right hand side of (16), that is, for (14). In order to do that we consider two cases.

1. If for any point P we have $\rho_P < \frac{\delta}{\omega}$, then

$$\begin{aligned}(1-\rho_P)^\omega + \omega\rho_P - 1 &= 1 - \omega\rho_P + \binom{\omega}{2}\rho_P^2 + \ldots + (-1)^\omega \binom{\omega}{\omega}\rho_P^\omega + \omega\rho_P - 1 \\ &\geq \rho_P^2\left(\binom{\omega}{2} - \rho_P\binom{\omega}{3} + \ldots + (-1)^\omega \rho_P^{\omega-2}\binom{\omega}{\omega}\right) \\ &\geqslant \rho_P^2\left(\binom{\omega}{2} - O_\omega(\delta)\right).\end{aligned}$$

   Hence (14) becomes

$$\frac{\sum_{\mathbb{P}} \kappa(\mathbb{P})}{|S|^\omega} \geq \frac{\omega(\omega-1)}{2} \sum_P \rho_P^2 - O_\omega\left(\delta \sum_P \rho_P^2\right). \tag{17}$$

   Using Cauchy's inequality for the inner sum in $O(\cdot)$

$$\sum_P \rho_P^2 \geqslant \frac{1}{\alpha|f|}\left(\sum_P \rho_P\right)^2 = \frac{1}{\alpha|f|}.$$

   Inserting it into (16) we get

$$\sum_{\mathbb{P}}^{*} \frac{\operatorname{ord}_f W(\mathbb{P})}{|S|^\omega} \geqslant \left(\frac{\omega(\omega-1)}{2\alpha} - O_{\omega,\alpha}(\delta)\right)\frac{1}{|f|} + O_{\omega,\alpha,|f|}(|S|^{-1}). \tag{18}$$

2. Suppose now that there is a point P, such that $\rho_P \geq \frac{\delta}{\omega}$. Using the fact that

$$\frac{\partial}{\partial x}((1-x)^\omega + \omega x - 1) = \omega(1-(1-x)^{\omega-1}) \geq \omega(1-(1-x)) = \omega x,$$

we get that

$$(1-\rho_P)^\omega + \omega\rho_P - 1 \geq \frac{1}{2}\omega\rho_P^2 \geq \frac{\delta^2}{2\omega}.$$

For P' different from P we have

$$(1-\rho_{P'})^\omega + \omega\rho_{P'} - 1 \geq 0$$

and $\sum_P \rho_P^2 \leq 1$. One can state

$$\begin{aligned}\frac{\operatorname{ord}_f W(\mathbb{P})}{|S|^\omega} &\geq \sum_P((1-\rho_P)^\omega + \omega\rho_P - 1) + O_\omega\left(|S|^{-1} + \delta\sum_P \rho_P^2\right)\\ &\geq \left(\frac{\omega}{2} + O_\omega(\delta)\right)\sum_P \rho_P^2 + O_\omega(|S|^{-1})\\ &\geq \left(\frac{\omega}{2} + O_\omega(\delta)\right)\frac{\delta^2}{\omega^2} + O_\omega(|S|^{-1}).\end{aligned}$$

For f with $|f| > c = c_{\omega,\alpha}$ the bound above implies (18), thus we use (18).

Multiply (18) by $\log_q|f|$ and sum over all f such that $\deg f | \operatorname{ord}_q N$ and $|f| > c$ (denoted by **)

$$\left(\frac{\omega(\omega-1)}{2\alpha} - O_{\omega,\alpha}(\delta)\right)S_1 + O_{\omega,\alpha,|f|}\left(\frac{S_2}{|S|}\right) \leqslant \sum_{\deg f|\operatorname{ord}_q N}^{**}\sum_{\mathbb{P}}^{*}\frac{\log_q|f|\operatorname{ord}_f W(\mathbb{P})}{|S|^\omega},$$

where

$$\begin{aligned}S_1 &= \sum_{\deg f|\operatorname{ord}_q N}^{**}\frac{\log_q|f|}{|f|},\\ S_2 &= \sum_{\deg f|\operatorname{ord}_q N}^{**}\log_q|f|.\end{aligned}$$

Since $\operatorname{ord}_f W(\mathbb{P}) \leqslant \log_{|f|}|W(\mathbb{P})|$, then $\log_q|f|\operatorname{ord}_f W(\mathbb{P}) \leqslant \log_q|W(\mathbb{P})|$. Thus applying (12) the expression above becomes

$$S_1\left(\frac{\omega(\omega-1)}{2\alpha} - O_{\omega,\alpha}(\delta)\right) + O_{\omega,\alpha,|f|}(S_2|S|^{-1}) \leqslant d_{\mathcal{W}}\log_q|I| + O_{\mathcal{W}}(1). \tag{19}$$

Further

$$S_2 = \sum_{f:\deg f|\operatorname{ord}_q N}\deg f = \sum_{i|\operatorname{ord}_q N} iM_i(q) = q^{\operatorname{ord}_q N} \leq N,$$

where $M_i(q)$ counts the number of irreducible polynomials of degree i and $M_i(q)$ has the property $\sum_{i|n} iM_i(q) = q^n$. Similarly we have $S_1 \geq \log_q N$. Inserting expressions for S_1 and S_2 in (19) and taking $N = |S|$ we get

$$\log_q|S|\left(\frac{\omega(\omega-1)}{2\alpha} - O_{\omega,\alpha}(\delta)\right) + O_{\omega,\alpha,|f|}(1) \leqslant d_{\mathcal{W}}\log_q|I| + O_{\mathcal{W}}(1).$$

Finally we end up with

$$|S| \ll_{\mathcal{W},c,\delta} |I|^{\frac{2\alpha d_{\mathcal{W}}}{\omega(\omega-1)}+O_{\omega,\alpha}(\delta)}$$

as demanded. □

Lemma 2. *Let $\mathcal{C}$ be an irreducible algebraic curve of degree d over $\mathbb{F}_q[T]$ which is defined by $F(X,Y)=0$. There exists a linear transformation*

$$(X,Y)\to(X',Y'),$$

such that $\deg_{X'} F(X',Y')=d$.

Proof. We can assume $\deg_X F(X,Y)<d$, otherwise we are done. Any polynomial of the form $F(X,Y)\in(\mathbb{F}_q[T])[X,Y]$ can be written as

$$F(X,Y)=\sum_{\substack{i\in J_1\\ j\in J_2}} F_{ij}X^iY^j,$$

where $J_1,J_2\subset\{0,1,...,d\}$, $F_{ij}\in\mathbb{F}_q$ and

$$\max_{\substack{i\in J_1\\ j\in J_2}}(i+j)=\deg F=d,$$
$$\max_{i\in J_1} i=\deg_X F<d.$$

Consider a linear transformation

$$(X,Y)\to(X',Y'),$$

such that

$$(X,Y)=(AX'+BY',CX'+DY'),$$

where $A,B,C,D\in\mathbb{F}_q[T]$ with $AD-BC\neq\mathbf{0}$. Changing the variables

$$(X,Y)\to(X',Y')$$

we obtain

$$\begin{aligned}F(X,Y)&=\sum_{\substack{i\in J_1\\ j\in J_2}} F_{ij}(AX'+BY')^i(CX'+DY')^j\\ &=\sum_{\substack{i\in J_1\\ j\in J_2}}\sum_{k=0}^{i}\sum_{l=0}^{j}\binom{i}{k}\binom{j}{l}F_{ij}A^{i-k}B^kC^{j-l}D^l(X')^{i+j-k-l}(Y')^{k+l}.\end{aligned}$$

In new variables (X',Y') we have

$$\deg_{X'} F=\max_{\substack{k\in\{0,...,i\},i\in J_1\\ l\in\{0,...,j\},l\in J_2}}(i+j-k-l),$$

which can be equal to d, since

$$\max_{\substack{i\in J_1\\ j\in J_2}}(i+j)=\deg F=d.$$

Denote by $F_d(X,Y)$ the homogeneous part of the highest degree. Then since $A,C\in\mathbb{F}_q[T]$, one can choose A and C, such that $F(A,C)\neq 0$. □

2.3 Proof of Theorem 1

Now we are ready to prove Theorem 1. We start with an interpolation argument, which is used for a similar goal in [21]. Let again $F \in (\mathbb{F}_q[T])[X,Y]$ be written in a form

$$F(X,Y) = \sum_{\substack{i\in J_1\\ j\in J_2}} F_{ij}X^iY^j,$$

where $J_1, J_2 \subset \{0,1,...,d\}$, $F_{ij} \in \mathbb{F}_q$. We are counting the number of distinct lattice points $P = (X,Y) \in I^2 \cap \mathcal{C}$. If we have less than $r(d) = d^2+1$ such points, then we are done. Suppose that we have at least $r(d)$ points: $P_i = (X_i, Y_i) \in C \cap I^2$, $i = 1,\dots,r(d)$ with $F(P_i) = \mathbf{0}$. Denote by $n(d) = \frac{1}{2}(d+1)(d+2)$ the number of monomials of degree less or equal than d. Consider $n(d) \times r(d)$ matrix A, whose i-th row consists of the monomials of degree d in the variables X_i, Y_i. Let $\vec{b} \in F_q^{n(d)}$ be a vector, whose entries are the corresponding coefficients F_{ij} of $F(X,Y)$. For such a vector $\vec{b}$ we have an equation

$$A\vec{b} = \vec{0}.$$

Since $\vec{b} \neq \vec{0}$, then the matrix A has a rank less than or equal to $n(d)-1$. Thus there is a solution $\vec{g} \neq \vec{0}$, where $\vec{g}$ is constructed out of the polynomials $X_i, Y_i \in I$, so we have $|\vec{g}| \ll_d |I|^{dn(d)}$. Let $G \in (\mathbb{F}_q[T])[X,Y]$ be the form of degree d corresponding to the vector $\vec{g}$. Then $G(X,Y)$ and $F(X,Y)$ share $r(d)$ zeros (points P_i). By Bézout's we know that since $F(X,Y) = G(X,Y) = \mathbf{0}$ contains more than $\deg F \deg G$ points, then F and G have a common factor. Since F is irreductible, F divides G and since $\deg G \leq \deg F$ we have $F = aG$. Let us work with G instead of F.

We are going to proceed in two steps:

1. If $\deg_X G < d$, then by Lemma 2 we can change variables so that $\deg_{X'} G = d$. If not, then proceed to the next step.

2. If f is such that $G(\mathrm{mod}\, f)$ is no longer irreducible (call these polynomials "bad"), then by Weil bounds on every irreducible component of G, say, G_i, we obtain

$$|\{(X,Y) \in (\mathbb{F}_q[T] \bmod f)^2 :\ G_i(X,Y) = \mathbf{0} \bmod f\}| \leq |f| + O_d(\sqrt{|f|}).$$

Such components satisfy

$$\prod_{f \text{ is bad}} |f| \leq |I|^{dn(d)}.$$

We subdivide all points of S according to which irreducible component of $G(\mathrm{mod}\, f)$ they reduce to modulo every bad f. The number of irreducible components of $G(\mathrm{mod}\, f)$ for every such f is less or equal than d. Then S is covered by sets $S_1, S_2, \dots, S_k$ with $k \leq d^{\#\{f:\ f \text{ is bad}\}} \ll_{d,\varepsilon} |I|^{\varepsilon}$. Every such set S_i intersects at most $|f| + O_d(\sqrt{|f|})$ residue classes modulo every irreducible $f \in \mathbb{F}_q[T]$. Further, for every $\varepsilon > 0$ and for every irreducible polynomial $f \in \mathbb{F}_q[T]$ with the condition $|f| \geqslant c(\varepsilon)$ the set S_i intersects at most $\left(1+\frac{\varepsilon}{2}\right)|f|$ residue classes mod f (here $c(\varepsilon)$ is a constant that depends only on ε). Applying Lemma 1 to S_i instead of S with $\alpha = 1+\frac{\varepsilon}{2}$ and $\mathcal{W}$ from Example 1: $\mathcal{W} = \mathcal{W}_{d-1,M}$ we obtain

$$|S| \ll_{\mathcal{W},\delta,c} |I|^{\frac{2\left(1+\frac{\varepsilon}{2}\right)d_{\mathcal{W}}}{\omega(\omega-1)}+O_{\omega,\varepsilon}(\delta)},$$

where (see Example 1)

$$d_{\mathcal{W}} = \sum_{F \in \mathcal{W}} \deg F = d(M+1)\frac{d+M-1}{2},$$
$$\omega = |\mathcal{W}| = d(M+1).$$

Thus

$$|S| \ll_{\varepsilon,d,\delta,M} |I|^{\frac{(1+\frac{\varepsilon}{2})(d+M-1)}{(d(M+1)-1)}+O_{\varepsilon,d,M}(\delta)}.$$

We choose $M = M(d)$ to be large enough to end up with

$$|S| \ll_{\varepsilon,d,\delta} |I|^{\frac{1}{d}+\frac{3\varepsilon}{4}+O_{\varepsilon,d}(\delta)}.$$

On taking δ to be small enough for $O_{\varepsilon,d}(\delta) \leq \frac{\varepsilon}{4}$ we obtain that

$$|S| < d(d+M-1)\delta^{-1} \ll_{d,\varepsilon} 1$$

and we are done.

2.4 An application to counting elliptic curves

In this section we are going to proceed with counting the number of elliptic curves $E_{a,b}$ with coefficients a, b living in a small box that lie in the same isomorphic classes. This is basically the generalization of several statements presented in [8]. Doing this we have an opportunity to apply Theorem 1 and also to show that some results for number fields can be also adapted to function fields.

Let I stand again for an interval of polynomials of the form $X(T)+Y(T)$, where $X(T) \in \mathbb{F}_q[T]$ is a fixed polynomial and $Y(T) \in \mathbb{F}_q[T]$ runs through all polynomials of degree less or equal than d. The coefficients of X and Y belong to $\mathbb{F}_q$ just as in section 2.

For a prime power q we consider a family of elliptic curves $E_{a,b}$

$$E_{a,b} : Y^2 = X^3 + aX + b,$$

where X and Y belong to $\mathbb{F}_q[T]$ as before and a, b are some coefficients from $\mathbb{F}_q[T]$ with the property that $4a^3+27b^2 \neq \mathbf{0}$. Let f be an irreducible polynomial over $\mathbb{F}_q[T]$. As in the number field case we say that two curves $E_{a,b}$ and $E_{c,d}$ are isomorphic over $\mathbb{F}_q[T]$ modulo f if there exists an invertible t modulo f such that

$$at^4 \equiv c (\mathrm{mod}\, f) \quad \text{and} \quad bt^6 \equiv d (\mathrm{mod}\, f).$$

The existence of an isomorphism between $E_{a,b}$ and $E_{c,d}$ implies that

$$a^3d^2 \equiv c^3b^2 (\mathrm{mod}\, f). \tag{1}$$

For $\lambda \in \mathbb{F}_q[T]$ we write $N_\lambda(I^2)$ for the number of solutions to the congruence

$$a^3 \equiv \lambda b^2 (\mathrm{mod}\, f), \quad (a,b) \in I^2.$$

We are going to give an upper bound on $N_\lambda(I^2)$ that implies upper bounds for the number of elliptic curves $E_{a,b}$ with coefficients $a, b \in I$ that lie in the same isomorphic classes.

For a polynomial $X \in \mathbb{F}_q[T]$ and an irreducible polynomial $f \in \mathbb{F}_q[T]$ we use $\{X\}_f$ to denote

$$\{X\}_f = \min_{Y \in \mathbb{F}_q[T]} |X - fY| = \min_{Y \in \mathbb{F}_q[T]} q^{\deg(X-fY)}.$$

From Dirichlet pigeon-hole principle we obtain

Lemma 3. *For real numbers $T_1, \dots, T_s$ with $1 \leqslant T_1, \dots, T_s \leqslant |f|$, $T_1 \cdots T_s \geqslant |f|^{s-1}$ and any polynomials $X_1, \dots, X_s \in \mathbb{F}_q[T]$ there exists a polynomial $t \in \mathbb{F}_q[T]$ such that t is not a multiple of f and*

$$\{X_i t\}_f \ll T_i, \quad i = 1, \dots, s.$$

Now we can give a good bound for $N_\lambda(I^2)$.

Theorem 1. *Let I be an interval of polynomials of degree less or equal than d with coefficients in $\mathbb{F}_q$ and the length of I is $|I| = q^d$. For any irreducible polynomial $f \in \mathbb{F}_q[T]$ such that $1 \leqslant |I| \leqslant |f|^{\frac{1}{9}}$ and for any $\lambda \in \mathbb{F}_q[T]$ we have*

$$N_\lambda(I^2) \leqslant |I|^{\frac{1}{3}+o(1)}.$$

Proof. We have to estimate the number of solutions to

$$(X + X_0)^3 \equiv \lambda(X_0 + Y)^2 \pmod f.$$

This congruence is equivalent to

$$X^3 + 3XX_0^2 + 3X^2X_0 - \lambda Y^2 - 2\lambda X_0 Y \equiv \lambda X_0^2 - X_0^3 \pmod f. \tag{20}$$

For any $T \leqslant |f|^{\frac{1}{4}}/|I|^{\frac{1}{2}}$ we can apply Lemma 3 to

$$X_1 = 1,\ X_2 = 3X_0,\ X_3 = 3X_0^2,\ X_4 = -\lambda,\ X_5 = -2\lambda X_0$$

and

$$T_1 = T^4|I|^2,\ T_2 = T_4 = \frac{|f|}{T|I|}, T_3 = T_5 = \frac{|f|}{T}$$

and find that there exists t with $|t| \leqslant T^4|I|^2$ such that

$$\{3X_0 t\}_f \leqslant \frac{|f|}{T|I|},\ \{3X_0^2 t\}_f \leqslant \frac{|f|}{T},\ \{\lambda t\}_f \leqslant \frac{|f|}{T|I|},\ \{2\lambda X_0 t\}_f \leqslant \frac{|f|}{T}.$$

For $i = 1, \dots, 5$ denote by f_i a polynomial which satisfies $f_i = X_i t$. Then multiply (20) by t leads us to the equality

$$f_1X^3 + f_2X^2 + f_3X + f_4Y^2 + f_5Y + f_6 = fZ, \tag{21}$$

where

$$|f_1| \leqslant T^4|I|^2,\ |f_2|, |f_4| \leqslant \frac{|f|}{T|I|},\ |f_3|, |f_5| \leqslant \frac{|f|}{T},\ |f_6| \leqslant \frac{|f|}{2}.$$

Since for $X, Y \in I$ we have $|X|, |Y| \leqslant |I|$, then the left hand side of (21) is bounded above by

$$T^4|I|^5 + \frac{4|f||I|}{T} + \frac{|f|}{2}.$$

Thus using strong triangle inequality

$$|Z| \ll \frac{T^4|I|^5}{|f|} + \frac{|I|}{T} + 1.$$

Choosing $T \approx \frac{|f|^{\frac{1}{5}}}{|I|^{\frac{4}{5}}}$ and applying the condition $1 \leqslant |I| \leqslant |f|^{\frac{1}{9}}$ we end with the bound

$$|Z| \ll \frac{|I|^{\frac{9}{5}}}{q^{\frac{1}{5}}} + 1 \ll 1.$$

□

Application of Theorem 1 to the family of curves E_{x^2,x^3} with $|x| \leqslant |I|^{\frac{1}{3}}$ shows that the result of Theorem 1 can not be improved. Thus in general we are not able to get any bound stronger than $N_\lambda(I^2) = O(|I|^{\frac{1}{3}})$.

3 Bounds on the number of integral points on elliptic curves

3.1 Introduction

Let K be the field of rational functions on an algebraic curve of genus g over the constant field k of characteristic 0. Denote by M_K the set of all places v of K. For a finite subset $S \subset M_K$ we denote by $\mathcal{O}_S$ the ring of S-integers of K. Consider a non-constant elliptic curve E given by a minimal Weierstrass equation

$$y^2 + a_1xy + a_3y = x^2 + a_2x^2 + a_4x + a_6,$$

where all the coefficients $a_1, a_2, a_3, a_4, a_6 \in \mathcal{O}_S$. The set of S-integral points of an elliptic curve E is

$$E(\mathcal{O}_S) = \{P \in E(K) : x(P), y(P) \in \mathcal{O}_S\}.$$

Our main goal here is to prove the following

Theorem 3. *Let q be a prime power, $\mathbb{F}_q[T]$ is the field of polynomials in a formal variable T with coefficients in a finite field $\mathbb{F}_q$ of order p. Let E be an elliptic curve over $\mathbb{F}_q[T]$ of a conductor N. Assume that the integral points on E are on minimal model. Then the number of integral points on E satisfies*

$$\#E(\mathbb{F}_q[T]) \leq \exp\left(c\frac{\deg N_E}{\log\deg N_E}\right),$$

where c is an absolute constant and N_E is the degree of the conductor of E.

Notice, that we work in the context where the analogue of Siegel's theorem is true (it is proven in [50]). In particular, if E is an ellipic curve over $\mathbb{F}_q[T]$ parametrized by $a, b \in \mathbb{F}_q$, then $E(\mathbb{F}_q(T)) = E(\mathbb{F}_q)$ and $\#E(\mathbb{F}_q[T]) \leq q + 1 + 2\sqrt{q}$. For a more general function field $\mathbb{F}_q(C)$ with ring of integers A we can have $E(A)$ infinite. Notice that if E is constant, i.e. defined over $\mathbb{F}_q$, then $E(\mathbb{F}_q(T)) = E(\mathbb{F}_q)$, therefore Siegel theorem holds in this case too. For the case of E being isotrivial (not defined over F_q and supersingular) Siegel theorem may be false.

The tools that allow us to proceed are that the necessary part of the famous Birch and Swinnerton-Dyer conjecture holds in the function field context (see [43], [31]), as well as the bounds for the analytic rank over a function field are known, thanks to the explicit formula given by Brumer in [5]. We also extend the technique of Helfgott [23](based on ideas of Silverman [40]) to obtain an upper bound for the number of integral points on E in terms of its algebraic rank. However, this brings us to results that do depend on the curve. To get rid of this dependence we have to work with the estimation of the sort $\#E(\mathbb{F}_q[T]) \ll c^{\operatorname{rank} E+m}$ more carefully (here m stands for the number of multiplicative places). Namely, we extend the method developed by Helfgott-Venkatesh in [23]. We optimize the size of c by applying sphere packing results.

The previously known bounds of such a type, namely Theorem 1, give us $\#E(\mathbb{Z}\cap I^2) \ll |I|^{\frac{1}{3}+\varepsilon}$, where we are restricted to counting integral points lying in a small box of size $|I|$. This result is analogous to Bombieri-Pila theorem [4], that gives the upper bound $\ll N^{\frac{1}{d}+\varepsilon}$, where d is the degree of a curve and is equal to 3 in the case

of elliptic curves, however the method of getting it is different and mainly based on the ideas of Helfgott-Venkatesh [23] and the interpolation part used by Heath-Brown [21]. Here we take the approach proposed by Helfgott in [23] and further developed by Helfgott-Venkatesh in [23], but it turns out that this way of doing things is closely related to the one used in [4].

The section is organized as follows. First we review some basic definitions that are going to be used throughout the paper as well as some important facts (see (4) and (5), also (6)) that are crucial in our proof. Then we prove several standard results regarding canonical height on an elliptic curve E. Based on this we show how to get a cheap, but useless bound for the number of points in $E(\mathbb{F}_q[T])$ of a bounded height (see Corollary 1, that has a dependency on the curve, which is undesirable). We introduce local heights $\lambda_v(\cdot)$ to get rid of this problem and prove lower bounds for $\lambda_v(\cdot)$ under some 'good' slicing, that will bring us to another bound for the canonical height, namely Lemma 5, that is proved in the spirit of [23, Proposition 3.4]. We also need a lower bound for the canonical height on E due to Silverman, see [40] and [39].

Further we prove the bound for the number of S-integral points on E in terms of algebraic rank of E using Lemmas announced before together with sphere packing results by Kabatyansky and Levenstein [27]. Finally, we prove the main result by taking an advantage of working in function fields, where Birch and Swinnerton-Dyer conjecture partly holds (see (4)) and apply the explicit formula for an analytic rank, given in the expression (5) by Brumer and further developed by the results of Grothendick, Deligne, etc.

3.2 Preliminaries and notations

We review some basic information about elliptic curves over a function field K. For more detailed information, the survey by Ulmer [46] is a perfect source. Let $k = \mathbb{F}_q$ be the finite field of cardinality q, with its characteristics $\mathrm{char}(k) = p$. We write K for the function field of a smooth, projective absolutely irreducible curve $\mathcal{C}$ over k. In what follows we consider $\mathcal{C} = \mathbb{P}^1$, thus $K = \mathbb{F}_q[T]$ is the field of polynomials in a formal variable T with coefficients lying in k. For $X \in K$ we denote by $|X|$ its norm: $|X| = q^{\deg X}$. We recall that an elliptic curve over K is a smooth, projective, absolutely irreducible algebraic curve of genus 1 over K with a K-rational point $\mathcal{O}$ that plays the role of identity element in the group $E(K)$ of K-rational points lying on E. This group is also called the Mordell-Weil group of E. The classical Mordell-Weil result was generalized to the case of function fields by Lang and Néron, namely $E(K)$ is a finitely generated abelian group. As a consequence of this result the torsion group $E(K)_{\mathrm{tors}}$ (i.e. the group of K-points on E of finite order) is finite and isomorphic to a group of the form

$$\mathbb{Z}/m\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z},$$

where m divides n and p does not divide m. Define an algebraic rank(E) of an elliptic curve E/K as the number of independent points of infinite order in $E(K)$, so to say the number of copies of $\mathbb{Z}$ in $E(K)$. Throughout the section we use the fact that the torsion group is unifromly bounded (see, for example, [46, Proposition 7.1], which gives a bound on $E_{tors}(K)$ is terms of the discriminant of a number field K). A strengtening of this result was proven in [38], where the analogous bound in terms of the degree of a field K was given.

An equivalent definition of an elliptic curve E/K can be given due to the Riemann-Roch theorem. An elliptic curve E/K can always be described as a projective plane curve of degree 3 with a (homogeneous) Weierstrass equation

$$y^2z + a_1xyz + a_3yz^2 = x^3 + a_2x^2z + a_4xz^2 + a_6z^3,$$

where all a_i belong to K. As usually, the origin is the point at infinity, namely $\mathcal{O} = [0 : 1 : 0]$. The condition of smoothness of E is equivalent to the fact that its discriminant Δ is not zero. The equation above can be also given in an affine form by going from projective coordinates (x, y, z) to affine coordinates $(x/z, y/z)$

$$y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x + a_6.$$

Let v be an equivalence class of valuations of K. Recall that a valuation on a field K is a generalization of the p-adic norm. Concretely, it is a function $|\cdot|_v$ from a field K to the real numbers $\mathbb{R}$ such that the following properties hold for all $x, y \in K$:

- $|x|_v \geq 0$, $|x| = 0$ if and only if $x = 0$;
- $|xy|_v = |x|_v \cdot |y|_v$;
- $|x|_v \leq 1$ implies $|1 + x|_v \leq C$ for some constant $C \geq 1$ independent of x.

Notice that if a valuation $|\cdot|_v$ satisfies the last condition above with $C = 2$, then it satisfies the triangle inequality

$$|x + y|_v \leq |x|_v + |y|_v$$

for all $x, y \in K$ and such a valuation is called archimedean. If the condition is satisfied with $C = 1$, then $|\cdot|_v$ satisfies the stronger ultrametric inequality:

$$|x + y|_v \leq \max(|x|_v, |y|_v)$$

for all $x, y \in K$ and we call this valuation non-archimedean. Here we work only with non-archimedean valuations.

For every v denote by $\mathcal{O}_{(v)}$ the ring of rational functions on $\mathcal{C}$ regular at v. In our case ($\mathcal{C} = \mathbb{P}^1$) the finite places correspond to monic irreducible polynomials $f \in K = \mathbb{F}_q[T]$. If such a place v corresponds to f, then

$$\mathcal{O}_{(v)} = \{g/h. \text{ s.t. } g, h \in K, \deg(g) < \deg(h)\}.$$

Assume that the degree of $v = \infty$ is 1. Write $\mathcal{M}_v \subset \mathcal{O}_{(v)}$ for the maximal ideal (its elements are the functions vanishing at v) and $\kappa_v = \mathcal{O}_{(v)}/\mathcal{M}_v$ for the residue field at v. Set $\deg(v) = [\kappa_v : k]$, $q_v = q^{\deg(v)}$ for the norm of v. Choose a minimal integral model for E in the Weierstrass form. Let $\bar{a}_i \in \kappa_v$ be the reductions of the coefficients at v and define the reduced curve E_v by

$$E_v : \ y^2 + \bar{a}_1xy + \bar{a}_3y = x^3 + \bar{a}_2x^2 + \bar{a}_4x + \bar{a}_6 \tag{2}$$

over the residue field κ_v. We say that E_v has

- a good reduction at v if E_v defines an elliptic curve over κ_v ($v \nmid \Delta$),

- a multiplicative (nodal) reduction at v if E_v has a node at v. If the tangent lines at the node are rational over the residue field κ_v, then we call this type of reduction split multiplicative. Otherwise non-split multiplicative.
- an additive (cuspidal) reduction at v if E_v has a cusp at v.

Notice that terms multiplicative and additive are used here to emphasize that the non-singular part of the reduced curve defined by $E_v^* = E_v/\{\text{singular point}\}$ is isomorphic to $\mathbb{G}_m$ (or $\mathbb{G}_m[\cdot]$ for the non-split case) and $\mathbb{G}_a$ respectively (here $\mathbb{G}_m$ stands for the multiplicative group, $\mathbb{G}_m[\cdot]$ for the twisted multiplicative group and $\mathbb{G}_a$ for the additive group). Elliptic curves, $\mathbb{G}_a$, $\mathbb{G}_m$ and $\mathbb{G}_m[\cdot]$ over K are the only irreducible algebraic curves over K having group structures given by regular maps.

The reduced curve E_v may be singular, but yet the non singular part of it, denoted by $\tilde{E}_v(K_v)$ forms a group. Moreover $E(K)$ admits the following filtration of abelian groups

$$E_1(K) \subset E_0(K) \subset E(K),$$

where

$$\begin{aligned} E_0(K) &= \{P \in E(K) : P_v \in \tilde{E}_v(K_v)\}, \\ E_1(K) &= \{P \in E(K) : P_v = O_v\} \end{aligned} \tag{22}$$

with P_v taken to be the image of $P \in E(K)$ under the reduction map $E(K) \to \tilde{E}_v(K_v)$.

A model for E given by E_v with its coefficients $\bar{a}_i \in \mathcal{O}_{(v)}$ is called integral at v. The minimal integral model at v is the model E_v with the valuation of the discriminant Δ of E being minimal. The local exponent n_v of the conductor at v is given by

$$n_v = \begin{cases} 0, & \text{if } E \text{ has good reduction at } v, \\ 1, & \text{if } E \text{ has multiplicative reduction at } v, \\ 2 + \delta_v, & \text{if } E \text{ has additive reduction at } v, \end{cases}$$

where δ_v is the wild ramification

$$\delta_v = \begin{cases} 0, & \text{if } p > 3, \\ \geqslant 0, & \text{if } p = 2, 3. \end{cases}$$

Thus n_v has the information about the ramification in the field extensions generated by the points of finite order in the group law of the elliptic curve E. The conductor of E/K is given by a product of prime ideals and associated exponents n_v. The (global) conductor of E is a divisor

$$N = \sum_v n_v[v].$$

The degree of the conductor is

$$\deg N = \sum_v n_v \deg v.$$

N is an effective divisor on $\mathbb{P}^1$ which is divisible only by the places v of bad reduction of E. The L-function of E is defined as the Euler product

$$L(E, s) = \prod_{v \nmid N}^{\text{good}} \left(1 - \frac{a_v}{q_v^s} + \frac{q_v}{{q_v}^{2s}}\right)^{-1} \times \prod_{v|N}^{\text{mult}} \left(1 - \frac{1}{q_v^s}\right)^{-1} \tag{3}$$

where "good" stands for "E has a good reduction at v", "mult" – for the case of either split multiplicative or non split multiplicative reduction at v and, finally, a_v is an integer defined as

$$a_v = \begin{cases} q_v + 1 - \#E_v(k_v), & \text{if } E \text{ has good reduction at } v, \\ \pm 1, & \text{if } E \text{ has multiplicative reduction at } v, \\ 0, & \text{if } E \text{ has additive reduction at } v. \end{cases}$$

($a_v = 1$ for the split multiplicative reduction and $a_v = -1$ for the non split multiplicative reduction). Due to the Hasse bound on a_v the first product of (3) converges absolutely for $\operatorname{Re} s > 3/2$ and admits a meromorphic continuation on $\mathbb{C}$. As usually we define an analytic rank of E/K as the order of vanishing of its L-function at $s = 1$

$$\operatorname{rank}_{an}(E) = \operatorname{ord}_{s=1} L(E, s).$$

We recall that an elliptic curve E/K is called constant if it can be defined by a Weierstrass equation with coefficients belong to k. It is called non-constant if it is not constant. Also E/K is called isotrivial if it becomes constant over some finite extension of K, otherwise – non-isotrivial.

Remark. *In the non-constant case of E Theorem 9.3 of [46] gives us an upper bound of a type* $\operatorname{rank}_{an} E \leqslant N$.

The famous conjecture of Birch and Swinnerton-Dyer connects the analytic behaviour of L-functions of elliptic curves with the group of K-rational points on E/K, in particular (among some other relations) it predicts that

$$\operatorname{rank}_{an}(E) \stackrel{?}{=} \operatorname{rank}(E).$$

While the original conjecture remains unsolved, much more is known in this context for the case of function fields.

Theorem (Tate [43], Milne [31])**.** *Let E be an elliptic curve over a function field K. Then*

$$\operatorname{rank} E \leq \operatorname{rank}_{an} E. \tag{4}$$

The usual technique for obtaining upper bounds of an analytic rank is using so-called explicit formula. We refer here to the result given by [5].

Theorem (Brumer [5])**.** *Let E be an elliptic curve over $\mathbb{F}_q[T]$. Then its analytic rank is bounded by*

$$\operatorname{rank}_{an} E \leq \frac{(b_E - 4)\log q}{2\log b_E} + O\left(\frac{b_E \log^2 q}{\sqrt{q}\log^2 b_E}\right), \tag{5}$$

where b_E is the degree of L-function as a polynomial in q^{-s}.

For the case of $\mathbb{F}_q[T]$ we have

$$b_E = n_E - 4,$$

where $n_E = \deg N$ and N is the conductor of an elliptic curve E/K. We note that if E has a additive reductions and m multiplicative reductions, then

$$n_E \leq 2a + m.$$

This result is interesting if and only if n_E is rather big, since the trivial bound for the rank is $n_E + 4g_X - 4$. We thus have

$$\operatorname{rank}_{an} E \leq \frac{(\deg N - 8)\log q}{2\log\deg N} + O\left(\frac{\deg N \log^2 q}{\sqrt{q}\log^2 \deg N}\right). \tag{6}$$

The easy bound is

$$\operatorname{rank} E \leq \operatorname{rank}_{an} E \leq b_E = n_E - 4.$$

If E is constant, then $\operatorname{rank} E = 0$.

3.3 Heights and its properties

In this section we are going to investigate some properties of height function on an elliptic curve E over a field K. The crucial fact here is that $|\hat{h} - \frac{1}{2}h_x|$ and $|\hat{h}^E - \frac{1}{3}h_y|$ are bounded on the set of all points of E. This allows us to give a lower bound for $\hat{h}^E(P)$ as well as to estimate the number of points with $\hat{h}^E < c_2$ under condition that E does not have any non torsion points P with $\hat{h}^E(P) > c_1$. However, this path leads us to a problem that the bound would depend on the curve. To avoid this difficulty we will use local heights as in [18] and establish the bound

$$\lambda_v(P - Q) \geq \min(\lambda_v(P), \lambda_v(Q))$$

that fails only in the case of bad reduction with which we will deal separately. We subdivide $E(K_v)$ into small enough number of slices, so that

$$\lambda_v(P - Q) \geq \min(\lambda_v(P), \lambda_v(Q))$$

still holds true on these slices with P, Q belong to the same slice (for more details see Lemma 4 and Lemma 2). Using that we prove that integral points we wish to count are far apart from each other in the Mordell-Weil lattice. Recall that any elliptic curve over K can be written in the following form

$$E:\ y^2 = f(x), \tag{7}$$

where $f(x) \in K$ is a cubic polynomial defined by Weierstrass equation. We say that $d \in K$ is square free if it has no factor of the form g^2 with $g \in K$ and $\deg g \geq 1$. For any $d \in K$ square free define a quadratic twist of E as

$$E_d:\ dy^2 = f(x). \tag{8}$$

Note that we restrict to the case of square free d, since if d has a squared factor, then by a change of variables in (8) one can find a curve E_d^* isomorphic to E_d. We write $\hat{h}^E$ for the canonical height on an elliptic curve E, and h_x, h_y for the height on E with respect to x and y:

$$\hat{h}^E((x,y)) = \lim_{n\to\infty} \frac{1}{n^2} h_x([n](x,y)), \tag{9}$$

where we use the notation $[n]P = \underbrace{P + \ldots + P}_{n \text{ times}}$ and

$$h_x((x,y)) = \begin{cases} 0, & \text{if } P = \mathcal{O}, \\ \log_q H(x), & \text{otherwise,} \end{cases}$$

$$h_y((x,y)) = \begin{cases} 0, & \text{if } P = \mathcal{O}, \\ \log_q H(y), & \text{otherwise.} \end{cases}$$

For any $x \in K$ define its norm by $|x| = q^{\deg x}$. We notice that $\hat{h}^E$ is defined on all points of $E(\bar{K})$ and $\hat{h}$ is a positive definite quadratic form on $E(\bar{K})$ as well as on $E(K)$(in the sense that it maps non-torsion elements to positive numbers).

For $x = x_0/x_1$ with $x_0, x_1 \in K$ not having as polynomials any common factor other than a constant polynomial in K (we encrypt this fact by $(x_0, x_1)_K = 1$), one can write $H(x) = \max(|x_0|, |x_1|)$. Let L be any algebraic field extension of $\mathbb{F}_q[T]$. Define $H(y)$ by

$$H(y) = (H_L(y))^{[L:K]^{-1}},\ H_L(y) = \prod_w \max(|y|_w^{n_w}, 1),$$

where $y \in L$, the product is taken over all places w of L, n_w stands for the degree of quotient field $L_w/K_w[T]$. For example, if $y = \frac{y_0}{y_1}$ with $y_0, y_1 \in K$, then $y \in \mathbb{F}_q(T)$ and for $L = \mathbb{F}_q(T)$

$$H(y) = H_L(y) = \max(|y_0|, |y_1|).$$

We list some important properties of the canonical height in the following lemma.

Lemma 1. *Let $f(x) \in K$ be a monic, cubuc polynomial of non-zero discriminant in (7). Let also d be a square-free polynomial $d \in K$ and $P = (x, y)$ be a K-point on the quadratic twist E_d of E. Let $P' = (x, d^{1/2}y)$ be a point on $E_1 = E$ associated to P. Then*

1. *$\hat{h}^{E_d}(P) = \hat{h}^E(P')$, where the canonical heights are defined on E_d and E, respectively.*
2. *The height h_y $(y \neq 0)$ is bounded on E, namely $h_y(P') \geq \frac{3}{8} \deg d$.*
3. *$\hat{h}^{E_d}(P) \geq \frac{1}{8} \deg d + c_f$, where c_f is a constant depending only on f.*

Proof. **1.** We do not put any change in the x-coordinate, so clearly $h_x(P) = h_x(P')$. For the sake of simplicity we consider the case of $\operatorname{char} k \neq 2, 3$. The proof goes analogously in the characteristics 2 and 3. Under this assumption we can write an equation of E in so-called short Weierstrass form (see, for example, Theorem 2.1 in [32])

$$E : y^2 = x^3 + ax + b,\ a, b \in K. \tag{23}$$

Then the duplication law on E is given by

$$[2]P = P + P = \left(\frac{(3x^2 + a)^2 - 8xy^2}{4y^2}, \frac{F_{a,b}(x)}{(2y)^3} \right), \tag{24}$$

where

$$F_{a,b}(x) = x^6 + 5ax^4 + 20bx^3 - 5a^2x^2 - 4abx - a^3 - 8b^2.$$

The short Weierstrass equation for the twisted curve E_d is given by the change of variables $(x, y) \to (dx, d^2y)$

$$E_d : y^2 = x^3 + ad^2x + bd^3.$$

Write $X(P)$ and $Y(P)$ for the coordinate functions of P. Then

$$X([2]P') = \frac{(3x^2 + a)^2 - 8dxy^2}{4dy^2},$$
$$X([2]P) = \frac{(3x^2 + a)^2 - 8dxy^2}{4y^2}.$$

Thus $X([2]P') = X((P+P)')$. Further,

$$Y([2]P) = \frac{F_{a,b}(x)}{(2y)^3},$$
$$Y([2]P') = \frac{F_{a,b}(x)}{d^{\frac{3}{2}}(2y)^3},$$

which shows that $Y([2]P') = Y((P+P)')$. We conclude that $(P+P)' = P'+P'$. Notice that here the addition is made on E_d on the left hand side and on E on the right hand side. Iterating this and using (9) we get

$$\hat{h}(P) = \frac{1}{2}\lim_{n\to\infty}\frac{h_x([2^n]P)}{2^{2n}} = \frac{1}{2}\lim_{n\to\infty}\frac{h_x([2^n]P')}{2^{2n}} = \hat{h}(P').$$

2. Write $y = \frac{y_0}{y_1}$ for $y_0, y_1 \in K$, such that they do not have any common factor $g \in K$ of a positive degree. For $a, b \in K$ we denote by $\langle a, b\rangle = \langle a, b\rangle_K$ the biggest common factor (in the sense that there is no other polynomial $g \in K$ of a bigger degree, such that g is a factor of both a and b) of polynomials a, b. We have $\langle y_0, y_1\rangle_K = 1$ and we call such polynomials coprime. If g is a monic irreducible polynomial, such that g is a factor of $\langle d, y_1^2\rangle$, then g^2 can not be a factor of $\langle d, y_1^2\rangle$ (by the fact that d is a square free polynomial), but it is a factor of y_1. Hence, if g is not a factor of $\langle d, y_1^2\rangle$, then write

$$\langle d, y_1^2\rangle = \frac{dy_1^2}{\{d, y_1^2\}},$$

where $\{d, y_1^2\}$ is a minimal polynomial that has both d and y_1^2 as factors. Then using the fact that y_0 and y_1 are taken to be coprime we conclude that g has a power -1 as a factor of

$$dy^2 = dy_0y_1^{-2} = d^2y_0^2\langle d, y_1^2\rangle^{-1}\{d, y_1^2\}^{-1}.$$

Recall that P lies on our curve E, so $dy^2 = f(x)$ and if g has a non-negative degree as a factor of x, then it also has a non-negative degree as a factor of dy^2. But if g has a negative degree as a factor of x, then its degree in dy^2 drops to ≤ -3 leaving us with a contradiction. Therefore we conclude that $|y_1| \geq \langle d, y_1^2\rangle^2$. Since $y \in K$ we can write by the definition of $H(y)$ and considering the Euclidean norm

$$\begin{aligned} H(y) &= \max\left(|y_0||d^{-1}\langle d, y_1^2\rangle|^{-\frac{1}{2}}, |y_1||\langle d, y_1^2\rangle|^{-\frac{1}{2}}\right) \\ &\geq \max\left(|y_0||d^{-1}\langle d, y_1^2\rangle|^{-\frac{1}{2}}, |\langle d, y_1^2\rangle|^{\frac{3}{2}}\right) \geq |d|^{\frac{3}{8}}, \end{aligned}$$

where we used the fact that max gets its minimal value when $|\langle d, y_1^2\rangle| = |d|^{\frac{1}{4}}$. Finally,

$$h_y(P) = \log H(P) \geq \frac{3}{8}\log q^{\deg d} = \frac{3}{8}\deg d.$$

3. It is a simple consequence of 1 and 2. If P' is a point on $E = E_1$, then, by 1 $\hat{h}^{E_d}(P) = \hat{h}^E(P')$. The difference $|\hat{h}^E - h_x^E|$ is bounded on E, thus by application of second part of 1 the result follows. □

Corollary 1. *Let E be an elliptic curve over K. If there are no non-torsion points $P \in E(K)$ of a canonical height $\hat{h}(P) > c_1$, then there are at most*

$$O\left(\left(1 + 2\sqrt{\frac{c_2}{c_1}}\right)^{\operatorname{rank} E}\right)$$

points in $E(K)$ of a canonical height $< c_2$.

Proof. Let's take our canonical height to the square of the Euclidean norm. There is one to one correspondence $f : K^{\operatorname{rank} E} \to K^{\operatorname{rank} E}$ such that $\hat{h}^E(\mathbb{P}) = |f(\mathbb{P})|^2$ for all vectors $\mathbb{P} \in K^{\operatorname{rank} E}$ of the length $\operatorname{rank} E$ with coordinates in K. Since $\hat{h}^E(P) > c_1$ for all non-zero $P \in K$, then we are equipped by $f(K^{\operatorname{rank} E})$ with a lattice L, such that for every element $l \in L$ different from 0 we have $|l| \geq c_1^{\frac{1}{2}}$. For every point $l \in L$ draw a sphere Sp_l centred at l of the radius $\frac{1}{2}c_1^{\frac{1}{2}}$, so that they do not overlap. Each of the spheres Sp_l is contained in the bigger one Sp with the radius $c_2^{\frac{1}{2}} + \frac{1}{2}c_1^{\frac{1}{2}}$ centred at the origin. By bounding the total volume of all spheres by $\operatorname{vol}(Sp) \leq (c_2^{\frac{1}{2}} + \frac{1}{2}c_1^{\frac{1}{2}})^{\operatorname{rank} E}$ and using the fact that $\#E_{tors}(K)$ is uniformly bounded we end the proof. □

The implied constants c_1, c_2 do not have any dependency on the twist, but depend on the curve. This would bring us to a problem once we want to bound the canonical height in terms of naive height (namely, we want something of the sort $h(P) \leq c_3$, where $h(P)$ is the naive height and c_3 is an absolute constant), because then the constant inside big O will change to $(1 + 2\sqrt{c_3/c_1})^{\operatorname{rank} E}$, where c_1 depends only on the curve, whilst c_3 depends on both the curve and c_2 (say, $c_2 = c_3 + O_E(1)$). To avoid this difficulty we have to exclude the hidden dependency by the method proposed in [23].

Recall that κ_v is the residue field at v and $d_v = \deg(v) = [\kappa_v : k]$. Let M_k be the set of places v on K. For each place $v \in K$, there exists a natural local height function λ_v such that the canonical height on E can be given in terms of λ_v

$$\hat{h}^E(P) = \frac{1}{[K : \mathbb{Q}]} \sum_{v \in M_K} d_v \lambda_v(P).$$

We say that an elliptic curve E over a non-archimedean local field K has potentially good reduction if it has a model with good reduction in some extension of K. Similarly, E has potentially multiplicative reduction if it does not have potentially good reduction.

Lemma 4. *Let E be an elliptic curve over a non-archimedean local field K_v with potentially good reduction. Let $P, Q \in E(K_v)$ be two distinct points. Then*

$$\lambda_v(P - Q) \geq \min(\lambda_v(P), \lambda_v(Q)).$$

Proof. Consider an extension L_w of K_v on which E has good reduction. Choose a Weierstrass equation for E over L_w such that $v(\Delta) = 0$. Then by [18, Proposition 2] we find that

$$\lambda_v(P) = \lambda_w(P) = \frac{1}{2} \max(\log |x(P)|_w, 0).$$

Since v is non-archimedean, then $|x+y|_v \leq \max(|x|_v, |y|_v)$ and the claim follows. □

The following lemma is the result of [24].

Lemma 2. *Let E be an elliptic curve over a non-archimedean local field K_v with potentially multiplicative reduction. Then for any $\varepsilon > 0$ small enough, there is a subdivision*

$$E(K_v) = W_{v,0} \cup W_{v,1} \cup \ldots \cup W_{v,d_v}$$

with $d_v \ll |\log \varepsilon|$, such that for any two distinct points $P, Q \in W_{v,0}$ we have

$$\lambda_v(P - Q) \geq \min(\lambda_v(P), \lambda_v(Q)), \quad \lambda_v(P_1), \lambda_v(P_2) \geq 0,$$

and for any two distinct points $P, Q \in W_{v,j}$, where $1 \le j \le d_v$ we have

$$\lambda_v(P-Q) \ge (1-\varepsilon)\max(\lambda_v(P), \lambda_v(Q)),$$
$$\lambda_v(P-Q) \ge (1-2\varepsilon)\max(\lambda_v(P), \lambda_v(Q)),$$

where the implied constant is absolute.

Proof. To proceed we have to introduce Tate curve. Notice that in the case of elliptic curve $E/\mathbb{C}$ we know that there is always a parametrization of $\mathbb{C}/\Lambda$ for some lattice $\Lambda \subset \mathbb{C}$. Tate curve plays a role of such a parametrization for an elliptic curve E if we work with p-adic field instead of $\mathbb{C}$. For $q_* \in K_v^*$ with $|q|_v < 1$ we define a Tate curve E_{q_*} by a Weierstrass equation

$$E_{q_*} : \ y^2 + xy = x^3 + a_4(q_*)x + a_6(q_*),$$

where

$$a_4 = -5\sum_n n^3 \frac{{q_*}^n}{1-{q_*}^n},$$
$$a_6 = -\sum_n \frac{7n^5+5n^3}{12}\frac{q^{*n}}{1-{q_*}^n}.$$

This is an elliptic curve that satisfies (see, for example [41, Chapter V])

$$\Delta = q_*\prod_{n\ge 1}(1-{q_*}^n)^{24},$$
$$j(E_{q_*}) = \frac{1}{q_*} + 744 + 196884q_* + \ldots = \frac{1}{q_*} + \sum_{n\ge 0} c_n {q_*}^n.$$

The elliptic curve E is isomorphic over $\bar{K_v}$ to a Tate curve E_{q_*} for some $q_* \in K_v^*$ satisfying $v(q_*) = -v(j(E))$. Let $v(t) = -\log|t|$ and $\alpha(t) = \frac{v(t)}{v(q_*)}$. Let L_w be the minimal extension of K_v on which E has split multiplicative reduction. For every $P \in E(L_w)$ we set

$$\lambda(P) = -\frac{1}{2}B_2(\alpha(P))\log|q_*|_w + \delta(P)\log|\pi|_w,$$

where π is a uniformizer of w, $B_2(t) = t^2 - t + \frac{1}{6}$ and $\delta(P) = 0$ if $P \notin L_w$ and otherwise it is the largest integer n such that $P \in E_n(L_w)$, where E_n is defined through the canonical filtration of E_w(see Lang [29, pp 68-70] or Gross-Silverman [18, p. 270])

$$\ldots \subset E_1(L_w) \subset E_0(L_w) \subset E(L_w).$$

We have $\delta(P_1 - P_2) \ge \min(\delta(P_1), \delta(P_2))$. Thus for P_1, P_2 such that $\alpha(P_1) = \alpha(P_2) = 0$ we have

$$\lambda(P_i) = -\frac{1}{12}\log|q|_w - \delta(P_i)\log|\pi|_w, \quad i = 1, 2$$

and further

$$\begin{aligned}\lambda(P_1 - P_2) &= -\frac{1}{2}B_2(\alpha(P_1-P_2))\log|q|_w - \delta(P_1-P_2)\log|\pi|_w \\ &\le -\frac{1}{2}B_2(0)\log|q|_w - \min(\delta(P_1),\delta(P_2))\log|\pi|_w \\ &\le -\frac{1}{12}\log|q|_w - \min(\delta(P_1),\delta(P_2))\log|\pi|_w \\ &\le \min(\lambda(P_1), \lambda(P_2)).\end{aligned}$$

Now we are are almost in the part of partitioning $E(K_v)$ into "good" slices. Define

$$W_{v,0} = \{P \in E(K_v) : \alpha(P) = 0\}.$$

It remains to subdivide the following part

$$\{P \in E(K_v) : \alpha(P) \neq 0\}.$$

We are going to exploit the properties of $B_2(t)$ as a quadratic form. Let

$$\left(0, \tfrac{1}{2}\right] = U_0 \cup U_1 \cup \ldots \cup U_m,$$

where m is the smallest integer such that $m \geq \log_{\frac{3}{2}} \frac{6}{\varepsilon}$ and U_i are defined as

$$\begin{aligned} U_0 &= \left(0, \tfrac{\varepsilon}{12}\right], \quad U_m = \left(\left(\tfrac{3}{2}\right)^{m-1} \tfrac{\varepsilon}{12}, \tfrac{1}{2}\right] \\ U_j &= \left(\left(\tfrac{3}{2}\right)^{j-1} \tfrac{\varepsilon}{12}, \left(\tfrac{3}{2}\right)^{j} \tfrac{\varepsilon}{12}\right], \quad 1 \leq j \leq m. \end{aligned}$$

Let $t_1, t_2 \in U_j$ for some index j and $t_1 \geq t_2$. For $j = 0$ since

$$t_1, t_2 \in \left(0, \tfrac{\varepsilon}{12}\right] \text{ and } t_1 - t_2 \leq \frac{\varepsilon}{12},$$

then by applying the fact that $B_2(t)$ is decreasing on $t \in U_0 \cup U_1 \cup \ldots \cup U_m$ we get

$$\begin{aligned} B_2(t_1 - t_2) \geq B_2\left(\frac{\varepsilon}{12}\right) &= \frac{\varepsilon^2}{144} - \frac{\varepsilon}{12} + \frac{1}{6} \geq \frac{1}{6} - \frac{\varepsilon}{6} + \frac{\varepsilon}{12} = (1-\varepsilon)B_2(0) + \frac{\varepsilon}{12} \\ &> (1-2\varepsilon)B_2(0) + \frac{\varepsilon}{12}. \end{aligned}$$

Now consider $j \geq 1$. Then letting

$$u = \left(\frac{3}{2}\right)^{j-1} \frac{\varepsilon}{12}$$

we have

$$\begin{aligned} & u \leq t_2 \leq t_1 \leq \frac{3u}{2}, \\ & t_1 - t_2 \leq \frac{u}{2}, \\ & B_2(t_1 - t_2) \geq B_2\left(\tfrac{u}{2}\right). \end{aligned}$$

Consider the simpler case of $B_2(u) \geq \frac{1}{12}$. Then

$$\begin{aligned} B_2(t_1 - t_2) \geq B_2\left(\tfrac{u}{2}\right) \geq B_2(u) &= (1-\varepsilon)B_2(u) + \varepsilon B_2(u) \geq (1-\varepsilon)B_2(u) + \frac{\varepsilon}{12} \\ &> (1-2\varepsilon)B_2(u) + \frac{\varepsilon}{12}. \end{aligned}$$

It remains to work out the case of $B_2(u) < \frac{1}{12}$. In this case $\frac{1}{11} < u < 1$. We want to show the bound of the same type as in the previous case, so we want to show that

$$B_2(\tfrac{u}{2}) - (1-\varepsilon)B_2(u) - \frac{\varepsilon}{12} \stackrel{?}{>} 0$$

Since $u > \frac{1}{11}$ and $B_2(u) < \frac{1}{12}$, we have

$$B_2(\tfrac{u}{2}) - B_2(u) = \frac{u^2}{4} - \frac{u}{2} + \frac{1}{6} - B_2(u) > \frac{1}{4 \cdot 121} - \frac{1}{2 \cdot 11} + \frac{1}{12} > \frac{1}{30}.$$

Also we know that $B_2(u) \geq B_2(\frac{1}{2}) = -\frac{1}{12}$. Inserting these two expressions and taking $\varepsilon < \frac{1}{5}$ we have

$$\begin{aligned} B_2(\tfrac{u}{2}) - (1-\varepsilon)B_2(u) - \frac{\varepsilon}{12} &= B_2(\tfrac{u}{2}) - B_2(u) + \varepsilon B_2(u) - \frac{\varepsilon}{12} \\ &> B_2(\tfrac{u}{2}) - B_2(u) - \frac{\varepsilon}{6} > \frac{1}{30} - \frac{\varepsilon}{6} \geq 0. \end{aligned}$$

Analogously for $\varepsilon < \frac{2}{15}$ we have

$$\begin{aligned} B_2(\tfrac{u}{2}) - (1-2\varepsilon)B_2(u) - \frac{\varepsilon}{12} &= B_2(\tfrac{u}{2}) - B_2(u) + 2\varepsilon B_2(u) - \frac{\varepsilon}{12} \\ &> B_2(\tfrac{u}{2}) - B_2(u) - \frac{\varepsilon}{4} > \frac{1}{30} - \frac{\varepsilon}{4} \geq 0. \end{aligned}$$

Finally combining results for $j = 0$ and $j \geq 1$ we have

$$\begin{aligned} B_2(t_1 - t_2) &\geq (1-\varepsilon) \max_{j=1,2} B_2(t_j) + \frac{\varepsilon}{12}, \\ B_2(t_1 - t_2) &\geq (1-2\varepsilon) \max_{j=1,2} B_2(t_j) + \frac{\varepsilon}{12} \end{aligned}$$

for all $t_1, t_2 \in U_j$, $0 \leq j \leq m$ with $t_1 \geq t_2$. For $0 \leq j \leq m$ define

$$\begin{aligned} W_{v,2j+1} &= \{P \in E(K_v) : \alpha(P) \in U_j\}, \\ W_{v,2j+2} &= \{P \in E(K_v) : \alpha(-P) \in U_j,\ \alpha(P) \neq \tfrac{1}{2}\}. \end{aligned}$$

Setting $d_v = 2m + 2$ ends the proof. □

Now we have to adapt [24, Proposition 3.4], that will serve us for as a bound for the canonical height that does not depend on the curve any longer. Here we assume that our two points are of the same reduction as well as that they fall into the same W-class, so we can apply Lemma 2.

Lemma 5. *Let E be an elliptic curve over K. Let S be a finite set of places of $K = \mathbb{F}_q[T]$, that includes all irreducible divisors of the discriminant Δ of E. Let P_1, P_2 be two distinct integral points on E that belong to the same set $W_{v,i}$ for any place v among the ones with potentially multiplicative reduction. Suppose that*

$$\sum_{v \in T} d_v |\lambda_v(P_1) - \lambda_v(P_2)| \leq \varepsilon \max_{j=1,2} \sum_{v \in T} d_v \lambda_v(P_j),$$

where $\varepsilon > 0$ sufficiently small and

$$T = \{v \in S : \lambda_v(P_1), \lambda_v(P_2) \geq 0\}.$$

Assume that P_1 and P_2 have the same reduction modulo I, where I is any ideal not divisible by irreducible elements of S. Then

$$\hat{h}(P_1 - P_2) \geq (1-2\varepsilon) \max(\hat{h}(P_1), \hat{h}(P_2)) + \frac{\log NI}{[K:L]}.$$

Proof. If v is a finite place of good reduction, then $\lambda_v(P) \geq 0$. Recall that S contains all places that divide the discriminant Δ of E. Then by definition of a canonical height through local heights we have

$$\begin{aligned}\hat{h}(P_1 - P_2) &\geq \sum_{v \in S} d_v \lambda_v(P_1 - P_2) + \sum_{v \notin S} d_v \lambda_v(P_1 - P_2) \\ &= \sum_{v \in S} d_v \lambda_v(P_1 - P_2) + \sum_{\substack{v \text{ finite} \\ v(I) > 0}} d_v \lambda_v(P_1 - P_2).\end{aligned}$$

We now subdivide our set S as $S = T \cup S/T$, where T is defined in the statement of the lemma. Let us consider two differences

$$\begin{aligned}\sigma_1 &= \sum_{v \in T} d_v \lambda_v(P_1 - P_2) - (1 - \varepsilon) \sum_{v \in T} d_v \min(\lambda_v(P_1), \lambda_v(P_2)), \\ \sigma_2 &= \sum_{v \in S/T} d_v \lambda_v(P_1 - P_2) - (1 - 2\varepsilon) \max_{j=1,2} \sum_{v \in S/T} d_v \lambda_v(P_j).\end{aligned}$$

The goal now is to show that these two quantities $\sigma_1, \sigma_2 \geq 0$. Once we are done it remains to consider only finite places v, such that $v(I) > 0$. We use the following notations $\sum^{\text{good}}, \sum^{0}, \sum^{j}$ denote that P_1, P_2 are of potentially good reduction, potentially multiplicative reduction and fall into $W_{v,0}$, potentially multiplicative reduction and fall into $W_{v,j}$ with $j > 0$ respectively. By Lemma 4 and Lemma 2 we have

$$\begin{aligned}\sigma_1 &\geq \sum_{v \in T}^{good,0} d_v \min_{j=1,2} \lambda_v(P_j) + (1 - \varepsilon) \sum_{v \in T}^{j} d_v \max_{j=1,2} \lambda_v(P_j) - (1 - \varepsilon) \sum_{v \in T} d_v \min_{j=1,2} \lambda_v(P_j) \\ &= \varepsilon \sum_{v \in T} d_v \min_{j=1,2} \lambda_v(P_j) - \varepsilon \sum_{v \in T}^{j} d_v \max_{j=1,2} \lambda_v(P_j) + \sum_{v \in T}^{j} d_v \big(\max_{j=1,2} \lambda_v(P_j) - \min_{j=1,2} \lambda_v(P_j) \big) \\ &\geq \varepsilon \sum_{v \in T} d_v \min_{j=1,2} \lambda_v(P_j) - \varepsilon \sum_{v \in T}^{j} d_v \max_{j=1,2} \lambda_v(P_j) \\ &= \varepsilon \sum_{v \in T}^{good,0} d_v \min_{j=1,2} \lambda_v(P_j) + \varepsilon \sum_{v \in T}^{j} d_v \big(\min_{j=1,2} \lambda_v(P_j) - \max_{j=1,2} \lambda_v(P_j) \big) \\ &= \varepsilon \sum_{v \in T}^{good,0} d_v \min_{j=1,2} \lambda_v(P_j) - \varepsilon \sum_{v \in T}^{good,0} d_v \big(\min_{j=1,2} \lambda_v(P_j) - \max_{j=1,2} \lambda_v(P_j) \big) \\ &+ \varepsilon \sum_{v \in T} d_v \big(\min_{j=1,2} \lambda_v(P_j) - \max_{j=1,2} \lambda_v(P_j) \big) \\ &= \varepsilon \sum_{v \in T}^{good,0} d_v \max_{j=1,2} \lambda_v(P_j) + \varepsilon \sum_{v \in T} d_v \big(\min_{j=1,2} \lambda_v(P_j) - \max_{j=1,2} \lambda_v(P_j) \big).\end{aligned}$$

Now we apply the assumption of our lemma and get

$$\begin{aligned}\sigma_1 &\geq \varepsilon \sum_{v \in T}^{good,0} d_v \max_{j=1,2} \lambda_v(P_j) - \varepsilon^2 \sum_{v \in T} d_v \max_{j=1,2} \lambda_v(P_j) \\ &= (\varepsilon - \varepsilon^2) \sum_{v \in T}^{good,0} d_v \max_{j=1,2} \lambda_v(P_j) - \varepsilon^2 \sum_{v \in T}^{j} d_v \max_{j=1,2} \lambda_v(P_j) \geq 0\end{aligned}$$

by choosing ε small enough. Applying the same condition again we get

$$\begin{aligned}\sum_{v\in T} d_v\lambda_v(P_1-P_2) &\geq (1-\varepsilon)\sum_{v\in T} d_v \max_{j=1,2}\lambda_v(P_j)\\ &\quad+(1-\varepsilon)\sum_{v\in T} d_v\big(\min_{j=1,2}\lambda_v(P_j)-\max_{j=1,2}\lambda_v(P_j)\big)\\ &\geq (1-\varepsilon)\sum_{v\in T} d_v \max_{j=1,2}\lambda_v(P_j)+\sum_{v\in T} d_v\big(\min_{j=1,2}\lambda_v(P_j)-\max_{j=1,2}\lambda_v(P_j)\big)\\ &\geq (1-\varepsilon)\sum_{v\in T} d_v \max_{j=1,2}\lambda_v(P_j)-\varepsilon\sum_{v\in T} d_v \max_{j=1,2}\lambda_v(P_j)\\ &\geq (1-2\varepsilon)\sum_{v\in T} d_v \max_{j=1,2}\lambda_v(P_j).\end{aligned}$$

Similarly for σ_2

$$\sigma_2=\sum_{v\in S/T}^{good} d_v\big(\min_{j=1,2}\lambda_v(P_j)-(1-2\varepsilon)\max_{j=1,2}\lambda_v(P_j)\big)>0$$

with ε being small enough. Combining estimates for σ_1,σ_2 and using the fact that

$$\sum_{v\in T} d_v \max_{j=1,2}\lambda_v(P_j)\geq \max_{j=1,2}\sum_{v\in T} d_v\lambda_v(P_j)$$

one can see that

$$\begin{aligned}\sum_{v\in S} d_v\lambda_v(P_1-P_2) &\geq (1-2\varepsilon)\sum_{v\in T} d_v \max_{j=1,2}\lambda_v(P_j)+(1-2\varepsilon)\max_{j=1,2}\sum_{v\in S/T} d_v\lambda_v(P_j)\\ &\geq (1-2\varepsilon)\max_{j=1,2}\sum_{v\in S} d_v\lambda_v(P_j).\end{aligned}$$

Since S contains all places that do divide the discriminant, then we have

$$\lambda_v(P)=\frac{1}{2}\log^+(|x(P)|_v)=0,\ \text{for } v\notin S.$$

Then

$$\hat{h}_K(P_1-P_2)\geq (1-2\varepsilon)\max_{j=1,2}\hat{h}_K(P_j)+\sum_{\substack{v \text{ finite}\\ v(I)>0}} d_v\lambda_v(P_1-P_2).$$

It remains to consider only finite places v, such that $v(I)>0$. Let $\mathfrak{p}_v$ be the corresponding prime ideal in O_K with its multiplicity n_v in I. By reduction modulo $\mathfrak{p}_v^{n_v}$ our point P_1-P_2 becomes an origin O. Then

$$v(x(P_1-P_2))\leq -2n_v$$

and

$$\lambda_v(P_1-P_2)\geq \frac{n_v}{e_v}\log p_v,$$

where e_v is the ramification degree of K_v and p_v is the rational irreducible element under v. Thus

$$\sum_{\substack{v \text{ finite}\\ v(I)>0}} d_v\lambda_v(P_1-P_2)=\log NI.$$

□

We are going to exploit Lemma 5 to give an upper bound on the number of S-integral points. In order to get a good constant C, that appeared before in Theorem 3 we are going to apply sphere packing. We first subdivide the set of integer points on E into "good slices" and then apply sphere packing bounds to each part separately. Here we use the remarkable result of Kabatiansky and Levenstein (see, for example, [27]).

Lemma 3 (Kabatiansky-Levenstein [27])**.** *Let $A(n,\theta)$ be the maximal number of points that can be arranged on the unit sphere of $\mathbb{R}^n$ such that the angle between P_1, O and P_2 for any two P_1, P_2 of them is no smaller than θ. Then for $0 < \theta < \frac{\pi}{2}$*

$$\frac{1}{n}\log_2 A(n,\theta) \le \frac{1+\sin\theta}{2\sin\theta}\log_2\frac{1+\sin\theta}{2\sin\theta} - \frac{1-\sin\theta}{2\sin\theta}\log_2\frac{1-\sin\theta}{2\sin\theta} + o(1),$$

where the convergence is uniform and explicit for θ within any closed subinterval of $\left(0,\frac{\pi}{2}\right)$. In particular, for $\theta = \frac{\pi}{3}$, we have

$$\frac{1}{n}\log_2 A(m,\theta) \le 0.40141\ldots$$

Lemma 6. *Let c_1, c_2 be two positive real numbers, $0 < \varepsilon < \frac{1}{2}$, n is a non-negative integer. For $\vec{X} = (X_i)_{1\le i\le n} \in \mathbb{F}_q^n[T]$ consider*

$$S = \{\vec{X} \in \mathbb{F}_q^n[T]\ c_1 \le |\vec{X}| \le c_2\},$$

where

$$|\vec{X}| = \sum_{i=1}^{n} |X_i| = \sum_{i=1}^{n} q^{\deg X_i}.$$

Then one can cover S with balls $B(\vec{Y}, \varepsilon|\vec{Y}|)$, where $\vec{Y} \in T \subset \mathbb{F}_q^n[T]$ with

$$\#T \le C^n \varepsilon^{-(n+1)} \left(1 + \log\frac{c_2}{c_1}\right).$$

The implied absolute constant C is explicit.

Proof. It is enough to show the covering by balls $B(\vec{Y}, 2\varepsilon|\vec{Y}|)$. We wish to slice S into a union of regions where $|\cdot|$ is almost constant, namely

$$T = \bigcup_{0\le m\le M} \frac{c_1\varepsilon(1+\varepsilon)^m}{n} T_m,$$

where

$$T_m = \{\vec{Y} \in \mathbb{F}_q^n[T] :\ \frac{n}{\varepsilon}(1-\varepsilon) \le |\vec{Y}| \le \frac{n}{\varepsilon}(1+\varepsilon)\},$$
$$M = \log_{1+\varepsilon}\log\frac{c_2}{c_1}.$$

Let $\vec{X} \in S$. Consider

$$m(\vec{X}) = \left\lfloor \log_{1+\varepsilon}\frac{|\vec{X}|}{c_1} \right\rfloor,$$
$$\vec{Z}(\vec{X}) = \left\lfloor \frac{n\vec{X}}{c_1\varepsilon(1+\varepsilon)^{m(\vec{X})}} \right\rfloor,$$

where $\lfloor\cdot\rfloor$ is the floor function. Define

$$\vec{Y} = \frac{c_1\varepsilon(1+\varepsilon)^m}{n}\vec{Z}(\vec{X}).$$

Then

$$|\vec{Z}(\vec{X})| \leq \frac{n|\vec{X}|}{c_1\varepsilon(1+\varepsilon)^{m(\vec{X})}}$$

and thus $|\vec{Y}| \leq |\vec{X}| < c_2$. Next,

$$|\vec{Z}(\vec{X})| \geq \frac{n|\vec{X}|}{c_1\varepsilon(1+\varepsilon)^{m(\vec{X})}} - 1$$

and thus

$$|\vec{Y}| \geq |\vec{X}| - \frac{c_1\varepsilon(1+\varepsilon)^{m(\vec{X})}}{n} \geq c_1 - \frac{c_1\varepsilon(1+\varepsilon)^{M)}}{n} \geq c_1 - \frac{c_2\varepsilon}{n} > c_1.$$

We have just shown that given an $\vec{X} \in S$ one can find a point $\vec{Y}$, that depends on $\vec{X}$ and lies in T. In addition $\vec{Y}$ has the following property

$$d(\vec{X},\vec{Y}) = |\vec{X} - \vec{Y}| \leq 2\varepsilon|\vec{Y}|,$$

where $d(\cdot,\cdot)$ is the associated metric. It remains to estimate the size of T

$$\#T \leq \left(1 + \log_{1+\varepsilon}\frac{c_2}{c_1}\right)\#T_m \leq \left(1 + \log_{1+\varepsilon}\frac{c_2}{c_1}\right)\frac{(n(1+\frac{1}{\varepsilon})+n)^n}{n!}.$$

The result follows after application of Stirling formula. □

We will need the following lower bound for a canonical height on E.

Lemma 7. *Let E be an elliptic curve over K. There is an absolute constant $0 < c < 1$ such that, for every non-torsion point $P \in E(K)$ we have the bound*

$$\hat{h}(P) > c^m \max\big(1, h(j(E))\big),$$

where m is the number of mulpiplicative places and $j(E)$ is as usual a j-invariant of E.

Proof. This Lemma is an analogous result to the ones in [40] and [26]. In fact, a stronger result was proven in [26], namely: $\hat{h}(P) \geq c\sigma_E h(E)$, where σ_E is the Szpiro ratio (it gives $\hat{h}(P) \geq c_1 h(E)$ when $j(E) \in \mathbb{F}_q(T)/\mathbb{F}_q(T^p)$). □

3.4 Bounding the number of S-integral points

In this section we prove the bound for the number of S-integer points on E/K of height less than h_0. Here t is a parameter to be optimized further. Then we are going to present a proof of the main result. It follows the way proposed in [23], [40], [18] and later improved in [24]. By embedding $E(K)/E(K)_{tors}$ into $E(K)\otimes_{\mathbb{Z}}\mathbb{R} \cong \mathbb{R}^{\operatorname{rank} E}$ we can take the canonical height on E to be squared Euclidean norm. The key idea consists of the fact that the points we are looking at have large distance between each other. Namely, by choosing a good division of the area into small symmetric slices we can say that any two points are separated by almost 60 degrees. Then the number of integral points on E is bounded above by $2^{\operatorname{rank} E}$ (this constant was later improved to $(1+\varepsilon)$ in [24]). It remains to apply Theorem 4 and (6) for getting the result.

Theorem 2. *Let E be an elliptic curve over K. Let also S be a finite set of places of K, including all irreducible divisors of the discriminant of E. Then, for any $h_0 \geq 1$ and every $0 \leq t \leq 1$, the number of S-integer points P of $E(K)$ with a canonical height $\hat{h}(P) \leq h_0$ is at most*

$$O\left(C^{|S|}\varepsilon^{-2(|S|+[K:L])}|S|^{[K:L]}(1+\log h_0)^2 e^{t[K:L]h_0+(\beta(t)+\varepsilon)\operatorname{rank} E}\right),$$

where C is an absolute constant and $\beta(t)$ is defined for $0 \leq t < 1$ by

$$\beta(t) = \frac{1+f(t)}{2f(t)}\log\frac{1+f(t)}{2f(t)} - \frac{1-f(t)}{2f(t)}\log\frac{1-f(t)}{2f(t)},$$
$$f(t) = \frac{\sqrt{(1+t)(3-t)}}{2}, \quad \beta(1) = 0.$$

Proof. Briefly speaking, we subdivide S-integer points on E denoted by $E(K,S)$ into points $(\bmod I)$ for I being a suitable ideal in O_K. Then Lemma 5 states that after some manipulations on this partition the points, that lie in the same class tend to be far away from each other in the Mordell-Weil lattice. Here we apply sphere packing bounds of Kabatiansky and Levenstein, namely Lemma 3 to each part separately. These sphere packing bounds will bring us to the term $e^{\beta(t)\operatorname{rank} E}$ on each part. Summation over all the classes gives rise to another term $e^{[K:L]h_0}$. We have only to take care of getting the right conditions to apply Lemma 5.

We firstly subdivide $E(K,S)$ into a very few slices to force any two points of the same slice have comparable canonical height. Consider a set

$$\{P \in E(K,S) : \hat{h}(P) \leq h_0\}.$$

We want to cover it by sets of the form

$$\{P \in E(K,S) : (1-\varepsilon)h_i \leq \hat{h}(P) \leq h_i\}.$$

By Lemma 7 it is enough to take $\ll \varepsilon^{-1}(\log h_0 + |S|)$ such sets. Then we are allowed to decrease the power of $(1+\log h_0)^2$ just to 1, only for the set of points

$$\{(1-\epsilon)h_0 \leq \hat{h}(P) \leq h_0\}.$$

Suppose first that $t \neq 0$. Let S' be the set of places below S. If

$$X = \max(\lceil e^{th_0}\rceil, |\bar{S}|^{1+\frac{1}{[K:L]}}),$$

then there is an irreducible polynomial f in L, such that $f \notin \bar{S}$ and $X \leq |f| \leq 2X$. The ideal I of $\mathcal{O}_K$ generated by f satisfies

$$\frac{\log N(I)}{[K:L]} \geq h_0 t, \quad N(I) \ll_{[K:L]} s^{[K:L]+1} e^{th_0[K:L]}.$$

The S-integer points of our curve $E(K)$ fall into no more than $O_{[K:L]}(N(I))$ classes under the reduction modulo the corresponding ideal I. Define R to be the set of all places of potentially multiplicative reduction. For any place $v \in R$ we subdivide the corresponding $E(K_v)$ into $n_v + 1$ subsets, where n_v is defined as in Lemma 2(we take $\frac{\varepsilon}{2}$ instead of ε). Consider arbitrary tuples of the form $(a_v)_{v\in R}$, $(b_v)_{v\in R}$, such that $0 \leq a_v \leq n_v$ and $b_v = 0, 1$. We define B as the set of non-torsion points

$P \in E(K,S)$, such that for each $v \in R$ we have that P falls into the corresponding W-class – $P \in W_{v,a_v}$ and that $\lambda_v(p) \geq 0$ is equivalent to $b_v = 1$. Now we bound the number of elements in

$$B_{h_0} = \{P \in B : (1-\varepsilon)h_0 \leq \hat{h}(P) \leq h_0\}.$$

The number of such sets B is bounded above by $c_0^2|\log\varepsilon|^{s+[K:L]_\varepsilon-2[K:L]}$, that brings us to the desired result. Define $M = (S-R) \cup \{v \in R : b_v = 1\}$ and a map $l(P) = (d_v\lambda_v(P))_{v\in M}$. For $v \in S - M$ we know that $\lambda_v(P) < 0$, so that one can apply Lemma 7 and get

$$|l(P)|_1 > [K:L]\kappa^s \max(1, h(j)).$$

Using [18, Proposition 3] we get the bound

$$\sum_{v\notin M} d_v\lambda_v(P) \geq -\frac{1}{24}h_k(j) - 3[K:L].$$

On combining that we obtain

$$|l(P)|_1 \leq [K:L](h_0 + 3 + h(j)/24)$$

for $P \in B_{h_0}$. By Lemma 2 we can cover $l(B_{h_0})$ by at most

$$O(c_1^s\varepsilon^{-(s+1)}\log(h_0+1))$$

balls $B(x, \frac{\varepsilon}{8}|x|_1)$ in the 1-norm. Take two points $P_1, P_2 \in B_{h_0}$ with $l(P_i) \in B(x, \frac{\varepsilon}{8}|x|_1)$ for $i = 1, 2$. We then have

$$|l(P_1) - l(P_2)|_1 \leq \frac{\varepsilon}{4}|x|_1 \leq \frac{\varepsilon}{2}\max_{j=1,2}|l(P_j)|_1.$$

If these points have the same reduction modulo I, then we apply Lemma 5 and get that

$$\hat{h}(P_1 - P_2) \geq (1-\varepsilon)\max_{j=1,2}\hat{h}(P_j) + \frac{\log N(I)}{[K:L]} \geq (1+t-\varepsilon)\max_{j=1,2}\hat{h}(P_j).$$

Now we embed the Mordell-Weil lattice modulo torsion into $\mathbb{R}^{\operatorname{rank} E}$ by taking $\hat{h}$ to be the square of the Euclidean height. Since all $\hat{h}(P_1)$, $\hat{h}(P_2)$, $\hat{h}(P_1-P_2)$ are positive, then the images of P_1, P_2, say, $Q_1, Q_2 \in \mathbb{R}^{\operatorname{rank} E}$ aree different from each other and from the origin, so that the angle between them is at least $\arccos\frac{1-t+O(\varepsilon)}{2}$. We now apply Lemma 3 and get that there are at most $e^{r(\beta(t)+O(\varepsilon))}O_{[K:L]}(1)$ points of B_{h_0} with an image in a given ball and with a prescribed reduction modulo I. Now we combine these results with the number of variants for I, the number of possible sets B and the number of balls to get the theorem. Notice, that in the case $t = 0$ one simply proceeds without I. □

The case $t = 0$ is the pure application of sphere-packing results of Lemma 3, while the case $t = 1$ is related to the corresponding result of Bombieri-Pila type, namely Theorem 1.

Corollary 1. *Let E be an elliptic curve over K defined by a Weierstrass equation with integer coefficients. Let S be a finite set of places of K, including all places dividing the discriminant of E. Then for every sufficiently small ε the number of S integral points on E/K is at most*

$$O_\varepsilon\left(C^s\varepsilon^{-2(s+1)}\left(\log|\Delta|+\log p\right)^2 e^{\operatorname{rank} E(\beta(0)+\varepsilon)}\right).$$

We are now ready to give a version of Theorem 2 with an optimized parameter t.

We need as well upper bound for the canonical height. Here we adapt the result of Pacheco [34]. There are known bounds over $\mathbb{Q}$, see, for example [19]. Also one fins good bounds in [26], but they work only in characteristic 0.

Lemma 8. *Let E be an elliptic curve over K defined by a Weierstrass equation $y^2=f(x)$. Let $\mathcal{O}_S$ be the ring of S-inetegers and $\mathcal{O}_S^*$ be the ring of S-units. Suppose that $f(X)\in\mathcal{O}_S$ and the discriminant $\Delta\in\mathcal{O}_S^*$, $p>2$. Define a set Ξ in the following way. Let $f(X)=(X-x_1)(X-x_2)(X-x_3)$ be the factorization of $f(X)$ in $\bar{K}[X]$. Let $P=(x_P,y_P)\in\mathcal{O}_S$. Define $\xi_i^2=X-x_i$, $i=1,2,3$. Let $L=K(x_1,x_2,x_3,\xi_1,\xi_2,\xi_3)$. For any permutation $\{i,l,m\}$ of $\{1,2,3\}$ define*

$$\Xi=\left\{\frac{(\xi_i-\xi_l)}{(\xi_i-\xi_m)},\frac{(\xi_i-\xi_l)}{(\xi_i+\xi_m)},\frac{(\xi_i+\xi_l)}{(\xi_i-\xi_m)},\frac{(\xi_i+\xi_l)}{(\xi_i+\xi_m)}\right\}.$$

Then for any $\eta\in\Xi$ we have

$$\hat{h}_L(\eta)\le 2p^e(2g_L-2+|S_L|),$$

where S_L is the set of places of L lying over S and g_L is the genus of L. Moreover, if $p>3$, then for any $P=(x_P,y_P)\in\mathcal{O}_S$ we have $\hat{h}_K(y_P^4/\Delta)\le 48p^e(2g-2+|S|)$.

Corollary 2. *Let E be an elliptic curve over a field K. Let S be a finite set of places of K, that contains all places dividing the discriminant of E. Let $\alpha(x)=\min(xt+\beta(t),0\le t\le 1)$, where β is as in Theorem 2. Let also $R=\max(1,\operatorname{rank}E(K))$. Then for every $h_0\ge 1$ and for every sufficiently small ε, the number of S-integral points on E over K, that have canonical height less or equal to h_0 is at most*

$$O_{\varepsilon,[K:L]}\left(C^{\#S}\varepsilon^{-2(\#S+[K:L])}\#S^{[K:L]}(1+\log h_0)^2 e^{R\alpha([K:L]h_0/R)+\varepsilon R}\right),$$

where C is an absolute constant.

Bounds on an algebraic rank

Here we get the desired bound for an algebraic rank and give a bound for the number of S integral points on E in terms of its conductor. Due to the results of the previous section we have

$$\begin{aligned}\#E(K)&\ll c^{\operatorname{rank}E}\le c^{\operatorname{rank}_{an}E}\\&\le\exp\left(\log c\left(\frac{(\deg N-8)\log q}{2\log\deg N}+O\left(\frac{\deg N\log^2 q}{\sqrt{q}\log^2\deg N}\right)\right)\right),\end{aligned}$$

where we used the fact that $\operatorname{rank}E\le\operatorname{rank}_{an}E$ as well as the explicit formula given in Theorem 5. We see that the term in $O(\cdot)$ is smaller than the main term, so we can simply rewrite

$$\#E(K)\ll c^{\operatorname{rank}E+m}\le\exp\left(c\frac{\deg N\log q}{\log\deg N}\right).$$

3.5 Comparison to Bombieri-Pila type bound

Let S be the set of all points of bad reduction of an elliptic curve E/K. Consider $h_0 > c\max(\deg\Delta, h(j))$, where Δ is the discriminant and j is the j-invariant of E/K for some constant c. The main contribution to Theorem 2 and, respectively, Corollary 2 is given by $e^{R\alpha(h_0/R)}$. The minimum in α is attained to the left of $t = 1$. Since $h_0 > c\deg\Delta$, then $\alpha(h_0/R) < (1-\delta_0)h_0/R$, where δ_0 positive and depending only on c. Thus for any $\delta_1 \leq \delta_0$ we obtain a bound

$$\#E(K,S) \ll e^{(1-\delta_1)h_0},$$

while Bombieri-Pila type result brings us to e^{h_0}, thus this method gives an improvement in the exponent and also improves the corresponding results from [25].

Another possible way to get this sort of bounds is using the work of Bhargava et al. on bounding the size of 2-torsion group, see [3]. The authors of [3] proved the first nontrivial bounds on the sizes of 2-torsion subgroups of the class groups of cubic and higher degree number fields. This is also an improvement on the bounds on the number of integral points given in [25]. They also gave a result for the function fields, see [3, Theorem 7.1].

4 A variant of Bombieri-Vinogradov theorem with explicit constants

4.1 Introduction

For integer number a and $q \geq 1$, let

$$\psi(x;q,a) = \sum_{\substack{n\leq x \\ n\equiv a \bmod q}} \Lambda(n),$$

where $\Lambda(n)$ is the von Mangoldt function. The Bombieri-Vinogradov theorem is an estimate for the error terms in the prime number theorem for arithmetic progressions averaged over all q up to $x^{1/2}$, or, rather almost all up to $x^{\frac{1}{2}}$.

Theorem (Bombieri-Vinogradov)**.** *Let A be a given positive number and $Q \leq x^{1/2}(\log x)^{-B}$ where $B = B(A)$, then*

$$\sum_{q\leq Q} \max_{2\leq y\leq x} \max_{\substack{a \\ (a,q)=1}} \left|\psi(y,q,a) - \frac{y}{\varphi(q)}\right| \ll_A \frac{x}{(\log x)^A}.$$

The implied constant in this theorem is not effective, since we have to take care of characters, associated with those q that have small prime factors. For an individual q Generalized Riemann hypothesis (GRH) provides

$$\psi(x;q,a) - \frac{\psi(x)}{\varphi(q)} = O(\sqrt{x}(\log x)^2),$$

thus Bombieri-Vinogradov theorem gives a result on average of the same magnitude as GRH.

We note that for Q sufficiently large the bound in Bombieri-Vinogradov theorem is trivial. Indeed, $1 \leq y \leq x$ and $q \leq x^{\frac{1}{2}}$ on applying Brun-Titchmarsh inequality we have

$$\psi(y;q,a) \leq \psi(x;q,a) = \sum_{\substack{n=p^k\leq x \\ n\equiv a(\bmod q)}} \log n \ll \frac{1}{\varphi(q)} \frac{\log x}{\log \frac{x}{q}} \ll \frac{x}{\varphi(q)}.$$

This brings us to the trivial bound in Bombieri-Vinogradov theorem

$$\sum_{q\leq Q} \max_{2\leq y\leq x} \max_{\substack{a \\ (a,q)=1}} \left|\psi(y;q,a) - \frac{y}{\varphi(q)}\right| \ll \sum_{q\leq Q} \frac{x}{\varphi(q)} \ll x(\log Q) \ll x(\log x). \tag{25}$$

The first result of this chapter is

Theorem 4 (Bombieri-Vinogradov theorem with explicit constants)**.** *Let $x \geq 4$, $1 \leq Q_1 \leq Q \leq x^{\frac{1}{2}}$. Let also $l(q)$ denote the least prime divisor of q. Define $F(x,Q,Q_1)$ by*

$$F(x,Q,Q_1) = \frac{14x}{Q_1} + 4x^{\frac{1}{2}}Q + 15x^{\frac{2}{3}}Q^{\frac{1}{2}} + 4x^{\frac{5}{6}} \log \frac{Q}{Q_1}.$$

Then

$$\sum_{\substack{q\leq Q \\ l(q)>Q_1}} \max_{2\leq y\leq x} \max_{\substack{a \\ (a,q)=1}} \left|\psi(y;q,a) - \frac{\psi(y)}{\varphi(q)}\right| < c_1 F(x,Q,Q_1)(\log x)^{\frac{7}{2}},$$

where

$$c_1 = \frac{5}{4}E_0c_0 + 1 = 42.140461\ldots,$$
$$E_0 = \prod_p \left(1 + \frac{1}{p(p-1)}\right) = 1.943596\ldots,$$
$$c_0 = (2A_0)^{\frac{1}{2}} \frac{2^5}{3^{\frac{3}{2}}\pi(\log 2)^2}\left(2 + \frac{\log(\log 2)}{\log\frac{4}{3}}\right) = 16.93375\ldots,$$
$$A_0 = \max_{x>0}\left(\frac{\psi(x)}{x}\right) = \frac{\psi(113)}{113} = 1.03883\ldots.$$

Here we reduce this power to $(\log x)^{\frac{7}{2}}$ (from $(\log x)^{\frac{9}{2}}$ in [1]) by applying an explicit version for an upper bound for

$$b_k = \sum_{\substack{d\le V\\ d|k}} \mu(d),$$

where $\mu(d)$ is Mobius function, V is a given number, namely the following Lemma of [22].

Lemma 5 (Helfgott [22])**.** *For V large enough we have*

$$\left|\frac{1}{Y}\sum_{k\le Y}|b_k|^2 - L\right| \le \frac{V^2}{Y} + C,$$

where $C = 0.000023$, $L = 0.440729$.

This Lemma (due to Helfgott [22]) is an explicit version of the sum considered in [13], where it was shown that

$$\sum_{d_1,d_2\le Y}\frac{\mu(d_1)\mu(d_2)}{\gcd(d_1,d_2)}$$

tends to a positive constant as $Y\to\infty$. It is also suggested without proving that L can be about 0.440729. What Helfgott proves is that $L = 0.440729 \pm 0.000023$.

Further we improve Theorem 4 using Lemmata 4 and 6.

Theorem 5 (Bombieri-Vinogradov theorem, effective version)**.** *Let $x \ge 4$, $1 \le Q_1 \le Q \le x^{\frac{1}{2}}$. Let also $l(q)$ be the smallest prime divisor of q. Then*

$$\sum_{\substack{q\le Q\\ l(q)>Q_1}} \max_{2\le y\le x}\max_{\substack{a\\(a,q)=1}}\left|\psi(y;q,a) - \frac{\psi(y)}{\varphi(q)}\right| \ll x^{\frac{1}{2}}Q(\log x)^2 + \frac{x}{Q_1}(\log x)^3 + x^{\frac{95}{96}}(\log x)^4.$$

The constant can be made explicit by using the results of [22] in Lemma 3.

We can also give an ineffective version that follows from Propostion 2.

Corollary 2. *Let A be a positive number and $Q \le \frac{x^{1/2}}{(\log x)^A}$. Then Proposition 2 gives us the following bound*

$$\sum_{q\le Q}\max_{2\le y\le x}\max_{\substack{a\\(a,q)=1}}\left|\psi(y,q,a) - \frac{y}{\varphi(q)}\right| \ll_A \frac{x}{(\log x)^{A-2}}.$$

The implied constant in Corollary 2 is ineffective.

The improvement here consists in having a factor of $(\log x)^2$, rather than $(\log x)^{\frac{5}{2}}$ or $(\log x)^3$. In order to prove Proposition 2 (Vaughan inequality, which is the main step in proving Theorem 5) we use the weighted version of Vaughan's identity (see Lemma 4) and an estimate due to Barban-Vehov [2] and Graham [17]. While Graham uses the Siegel-Walfisz theorem, there is an effective (and explicit) version of the estimate in [22].

Proposition 2 allows us to prove Bombieri-Vinogradov theorem. In both effective (Theorem 5) and ineffective (Corollary 2). The proof of Theorem 2 uses the Siegel-Walfisz theorem, which states that

$$\psi(x,\chi) - \delta(\chi)x \ll_A xe^{-c\sqrt{\log x}}$$

uniformly for $q \leq (\log x)^A$. Here $A > 0$ is a fixed real number, c is an absolute positive constant, and $\delta(\chi) = 1$ if χ is principal and is zero otherwise. The implied constant in the Bombieri-Vinogradov theorem is ineffective since the implied constant in the Siegel-Walfisz theorem is ineffective. To prove Corollary 2 we use the Siegel-Walfisz theorem to deal with moduli $q \leq Q$ having small prime divisors and Proposition 2 to deal with the sum over the remaining moduli.

Notice that getting instead of $(\log x)^2$ in Theorem 5 something less (for example $(\log x)^2(\log\log x)^{-2}$), would imply the fully effective version of Bombieri-Vinogradov theorem. That is ultimately due to the fact that we will no longer need Siegel-Walfisz theorem. Instead once can use Landau-Page reuslt (see [35], [28] and also [44, Theorem 8.25])

Previously, the best result in the literature followed from [13]; it had $A - 5/2$ instead of $A - 2$. While [13] does not state the result in full – focusing on estimating a crucial sum – a complete form can be found in Theorem 4 (together with a fully explicit version). Another effective variant without explicit constants is given by Lenstra and Pomerance [30, Lemma 11.2] (with bigger power of log) in their work on Gaussian periods.

Remark. *Define*

$$\pi(x) = \sum_{p \leq x} 1 \qquad \text{and} \qquad \pi(x;q,a) = \sum_{\substack{p \leq x \\ p \equiv a \pmod q}} 1.$$

Then Theorem 5 under the same assumptions can be also formulated for $\pi(x)$, $\pi(x;q,a)$*:*

$$\sum_{\substack{q \leq Q \\ l(q) > Q_1}} \max_{2 \leq y \leq x} \max_{\substack{a \\ (a,q)=1}} \left| \pi(y;q,a) - \frac{\pi(y)}{\varphi(q)} \right| \ll x^{\frac{1}{2}} Q (\log x)^2 + \frac{x}{Q_1}(\log x)^3 + x^{\frac{95}{96}}(\log x)^4.$$

The proof of the remark is exactly the same as in [1], we just have to change the power of $(\log x)$.

4.2 Proof of Theorem 4

The key tool for the proof of Theorem 4 is Vaughan's identity, which we have to get in an explicit version for our goal. Define

$$\psi(y,\chi) = \sum_{n \leq y} \Lambda(n)\chi(n),$$

the twisted summatory function for the von Mangoldt function Λ and a Dirichlet character χ modulo q. We prove

Proposition 1 (Vaughan's inequality in an explicit form)**.** *For $x \geq 4$*

$$\sum_{q\leq Q} \frac{q}{\varphi(q)} \sideset{}{^*}\sum_{\chi(q)} \max_{y\leq x} |\psi(y,\chi)| < c_0(7x + 2Q^2 x^{\frac{1}{2}} + 5Q^{\frac{3}{2}} x^{\frac{2}{3}} + 4Qx^{\frac{5}{6}})(\log x)^{\frac{5}{2}},$$

where Q is any positive real number and $\sum^*_{\chi(q)}$ means a sum over all primitive characters $\chi(\mathrm{mod}\, q)$ and $c_0 = 16.93375...$ (as in Theorem 4).

The goal is to get an explicit version of $f(x,Q)$ by applying an improved version of Pólya-Vinogradov inequality (see [37] and [15] for improvements), that will reduce the coefficients of $f(x,Q)$ and then we can apply Lemma 5.

4.3 Vaughan inequality

Fix arbitrary real numbers $Q > 0$ and $x \geq 4$. In this section, we shall establish Proposition 1, which is the main ingredient in the proof of Theorem 4. Here we follow the ideas of [1] and applying the results from [22]. The main tool in the proof is the large sieve inequality (see, for example [33, p.561])

$$\sum_{q\leq Q} \frac{q}{\varphi(q)} \sideset{}{^*}\sum_{\chi(q)} \left| \sum_{m=m_0+1}^{m_0+M} a_m \chi(m) \right|^2 \leq (M+Q^2) \sum_{m=m_0+1}^{m_0+M} |a_m|^2, \tag{10}$$

from which it follows (see [1, Lemma 6.1] and also [9, Chapter 8.3]) that

$$\sum_{q\leq Q} \frac{q}{\varphi(q)} \sideset{}{^*}\sum_{\chi(q)} \max_{y} \left| \sum_{m=m_0}^{M} \sum_{\substack{n=n_0 \\ mn\leq y}}^{N} a_m b_n \chi(mn) \right| \leq$$

$$c_3(M'+Q^2)^{\frac{1}{2}}(N'+Q^2)^{\frac{1}{2}} \left(\sum_{m=m_0}^{M} |a_m|^2 \right)^{\frac{1}{2}} \left(\sum_{n=n_0}^{N} |b_n|^2 \right)^{\frac{1}{2}} L(M,N), \tag{11}$$

where $c_3 = 2.64...$, $L(M,N) = \log(2MN)$ and

$$M' = M - m_0 + 1,$$
$$N' = N - n_0 + 1$$

are the number of terms in the sums over m and n respectively. Here the a_m, b_n are arbitrary complex numbers.

4.4 Sieving and Vaughan's identity

We reduce to the case $2 \leq Q \leq x^{1/2}$. If $Q < 1$, then the sum on the left-hand side of (1) is empty and we are done. Next, $1 \leq Q < 2$ then only the $q = 1$ term exists and we have

$$\sum_{q\leq Q} \frac{q}{\varphi(q)} \sideset{}{^*}\sum_{\chi(q)} \max_{y\leq x} |\psi(y,\chi)| = \max_{y\leq x} \left| \sum_{n\leq y} \Lambda(n) \right| = \psi(x) \leq A_0 x, \tag{26}$$

which is better than the theorem. Finally, if $Q > x^{1/2}$, Theorem 1 follows from (11) with $M = m_0 = n_0 = 1$, $N = \lfloor x \rfloor$, $a_m = 1$, $b_n = \Lambda(n)$ by the estimate

$$\sum_{n \le x} \Lambda(n)^2 \le \psi(x) \log x \le A_0 x \log x.$$

From now on we assume $2 \le Q \le x^{1/2}$. Notice that the fact that we can restrict ourselves to the range $2 \le Q \le x^{1/2}$ allows us to apply Lemma 5 (otherwise it would make less sense, since the main term in Lemma 5 would be smaller than O^*-term). As in [1] we will use Vaughan's identity

$$\Lambda(n) = \lambda_0(n) + \lambda_1(n) + \lambda_2(n) + \lambda_3(n),$$

where

$$\begin{aligned}
\lambda_0(n) &= \begin{cases} \Lambda(n), & \text{if } n \le U, \\ 0, & \text{if } n > U, \end{cases} \\
\lambda_1(n) &= \sum_{\substack{hd=n \\ d \le V}} \mu(d) \log h, \\
\lambda_2(n) &= - \sum_{\substack{mdr=n \\ m \le U, d \le V}} \Lambda(m) \mu(d), \\
\lambda_3(n) &= - \sum_{\substack{mk=n \\ m > U, k > V}} \Lambda(m) \sum_{\substack{d \mid k \\ d \le V}} \mu(d).
\end{aligned}$$

Assume $y \le x$, $q \le Q$, and χ is a character mod q. We use the above decomposition to write

$$\psi(y, \chi) = s_0 + s_1 + s_2 + s_3,$$

where

$$s_i = \sum_{n \le y} \lambda_i(n) \chi(n).$$

Let U, V be non-negative functions of x and Q to be set later and denote the contributions to our main sum by

$$S_i = \sum_{q \le Q} \frac{q}{\varphi(q)} \sideset{}{^*}\sum_{\chi(q)} \max_{y \le x} |S_i|.$$

Easily we obtain

$$\sum_{q \le Q} \frac{q}{\varphi(q)} \sideset{}{^*}\sum_{\chi(q)} \max_{y \le x} |\psi(y, \chi)| \le S_0 + S_1 + S_2 + S_3.$$

The heart of the proof of Theorem 1.3 in [1] are the following estimates:

Lemma (Akbary, Hambrook [1]). *We have*

$$
\begin{aligned}
&S_0 \le A_0 U Q^2,\\
&S_1 < \left(x + Q^{\frac{5}{2}} V\right) (\log xV)^2,\\
&S_2 < S_2' + S_2'',\\
&S_2' < (x + Q^{\frac{5}{2}} U)(\log xU)^2,\\
&S_2'' < \frac{c_3}{\log 2}\left(x + Qx^{\frac{1}{2}}U^{\frac{1}{2}}V^{\frac{1}{2}} + 2^{\frac{1}{2}}QxU^{-\frac{1}{2}} + Q^2x^{\frac{1}{2}}\right)(\log 2UV)^2(\log 4x),\\
&S_3 < \frac{2^{\frac{3}{2}}A_1^{\frac{1}{2}}c_3}{\log 2}(x + QxV^{-\frac{1}{2}} + 2^{\frac{1}{2}}QxU^{-\frac{1}{2}} + Q^2x^{\frac{1}{2}})\left(\log\frac{2x}{V}\right)^{\frac{3}{2}}(\log e^3V)(\log 4x),
\end{aligned}
$$

where

$$
c_3 = \frac{2}{\pi}\left(\frac{2 + \log\log 2/\log\frac{4}{3}}{\log 2}\right).
$$

We estimate S_3 contribution with the use of Lemma 5. Writing s_3 as a dyadic sum we have

$$
s_3 = -\sum_{\substack{M=2^\alpha\\ \frac{1}{2}U<M\le x/V}}\ \sum_{\substack{U<m\le x/V\\ M<m\le 2M}}\ \sum_{\substack{V<k\le x/M\\ mk\le y}} \Lambda(m)\left(\sum_{\substack{d\mid k\\ d\le V}}\mu(d)\right)\chi(mk).
$$

Using the triangle inequality

$$
S_3 \le \sum_{\substack{M=2^\alpha\\ \frac{1}{2}U<M\le x/V}}\ \sum_{q\le Q}\frac{q}{\varphi(q)}\sideset{}{^*}\sum_{\chi(q)}\max_{y\le x}\left|\sum_{\substack{U<m\le x/V\\ M<m\le 2M}}\ \sum_{\substack{V<k\le x/M\\ mk\le y}} a_m b_k \chi(mk)\right|,
$$

where $a_m = \Lambda(m)$, and, as it was defined in the introduction $b_k = \sum_{d\mid k,\, d\le V}\mu(d)$. Now we apply the large sieve inequality (11) to get

$$
S_3 \le c_3 \sum_{\substack{M=2^\alpha\\ \frac{1}{2}U<M\le x/V}} (M' + Q^2)^{\frac{1}{2}}(K' + Q^2)^{\frac{1}{2}}\sigma_1(M)^{\frac{1}{2}}\sigma_2(M)^{\frac{1}{2}}L(M)
$$

where

$$
\begin{aligned}
&\sigma_1(M) = \sum_{V<k\le x/M}|b_k|^2, \quad \sigma_2(M) = \sum_{\substack{U<m\le x/V\\ M<m\le 2M}}|a_m|^2,\\
&L(M) = \log\left(\frac{2x}{M}\min\left(\frac{x}{V}, 2M\right)\right) \le \log 4x,
\end{aligned}
$$

where M' and K' denote the number of terms in the sums over m and k, respectively. From the definition of M' and N' we conclude

$$
\begin{aligned}
&M' = \min\left(2M, \frac{x}{V}\right) - \max\left(M+1, U+1\right) \le M,\\
&K' = \frac{x}{M} - (V+1) + 1 \le \frac{x}{M}.
\end{aligned}
$$

By Chebyshev estimate we have an upper bound

$$\sigma_2(M) \le \sum_{m\le 2M} \Lambda(m)^2 \le \psi(2M)\log 2M \le 2A_0 M\log 2M.$$

Thus by Cauchy inequality

$$S_3 \le c_3(\log 4x) \sum_{\substack{M=2^\alpha \\ \frac12 U<M\le x/V}} (M+Q^2)^{\frac12}\left(\frac{x}{M}+Q^2\right)^{\frac12}(2A_0M\log 2M)^{\frac12}\sigma_1(M)^{\frac12}. \quad (27)$$

Further

$$M(M+Q^2)\left(\frac{x}{M}+Q^2\right) = Mx+Q^2x+M^2Q^2+MQ^4$$

and

$$(\log 2M)^{\frac12} \le \left(\log\frac{2x}{V}\right)^{\frac12}.$$

Using Lemma 5 we get

$$(\sigma_1(M))^{\frac12} \le \frac{x}{M}(L+C) - V(L+C) + 2V^2,$$

that implies

$$S_3 \le c_3(2A_0)^{\frac12}(x+2^{\frac12}Q^{\frac12}xU^{-\frac12}+QxV^{-\frac12}+Q^2x^{\frac12})(\log 4x)\left(\log\frac{2x}{V}\right)^{\frac12} \sum_{\substack{M=2^\alpha \\ \frac12 U<M\le x/V}} 1.$$

Since

$$\sum_{\substack{M=2^\alpha \\ \frac12 U<M\le x/V}} 1 \le \frac{\log\frac{2x}{V}}{\log 2},$$

then

$$S_3 \le \frac{c_3}{\log 2}(2A_0)^{\frac12}(x+2^{\frac12}Q^{\frac12}xU^{-\frac12}+QxV^{-\frac12}+Q^2x^{\frac12})(\log 4x)\left(\log\frac{2x}{V}\right)^{\frac32}.$$

Combining it with results of Lemma 4.4 we get

$$S = \sum_{q\le Q}\frac{q}{\varphi(q)}\sum\nolimits^{*}_{\chi(q)} \max_{y\le x}|\psi(y,\chi)| \le c_4\,\mathrm{R}(x,Q,U,V)\,\mathrm{G}(x,V,U), \quad (28)$$

where

$$c_4 = \max\left\{A_0, \frac{c_3}{\log 2}, \frac{c_3}{\log 2}(2A_0)^{\frac12}\right\} = \frac{c_3}{\log 2}(2A_0)^{\frac12},$$

$$\begin{aligned}\mathrm{R}(x,Q,U,V) &= 4x + 2Q^2x^{\frac12} + UQ^2 + Q^{\frac52}(U+V)\\ &\quad + 2^{\frac12}Q^{\frac12}xU^{-\frac12} + 2^{\frac12}QxU^{-\frac12} + Qx^{\frac12}U^{\frac12}V^{\frac12} + QxV^{-\frac12},\end{aligned}$$

$$\mathrm{G}(x,V,U) = \max\left\{(\log xV)^2, (\log xU)^2, (\log 2UV)^2\log 4x, \left(\log\frac{2x}{V}\right)^{\frac32}\log 4x\right\},$$

Now let's specify U and V. If $x^{\frac{1}{3}} \le Q \le x^{\frac{1}{2}}$, then $U = V = x^{\frac{2}{3}}Q^{-1}$. Then putting that into previous expression we get for the factor

$$\begin{aligned}\mathrm{R}_1(x,Q) &= 4x + 2Q^2x^{\frac{1}{2}} + Qx^{\frac{2}{3}}(1+2^{\frac{1}{2}}) + Q^{\frac{3}{2}}x^{\frac{2}{3}}(2+2^{\frac{1}{2}}+1) + x^{\frac{7}{6}} \\ &\le 4x + 2Q^2x^{\frac{1}{2}} + 2Qx^{\frac{5}{6}} + Q^{\frac{3}{2}}x^{\frac{2}{3}}(2+2^{\frac{1}{2}}+1).\end{aligned}$$

where we used the fact that $x^{\frac{7}{6}} \le Qx^{\frac{5}{6}}$ and $Qx^{\frac{2}{3}} \le Qx^{\frac{5}{6}}$. Working in the same manner with G and keeping in mind the condition $x \ge 4$ we find that

$$\mathrm{G}_1(x,V,U) \le \left(\frac{4}{3}\log x\right)^{\frac{3}{2}} 2\log x = \frac{2^4}{3^{\frac{3}{2}}}(\log x)^{\frac{5}{2}}.$$

If $Q \le x^{\frac{1}{3}}$, we let $U = V = x^{\frac{1}{3}}$ and get

$$\begin{aligned}\mathrm{R}_2(x,Q) &= 4x + 2Q^2x^{\frac{1}{2}} + Q^2x^{\frac{1}{3}} + 2Q^{\frac{5}{2}}x^{\frac{1}{3}} + 2^{\frac{1}{2}}Q^{\frac{1}{2}}x^{\frac{5}{6}} + Qx^{\frac{5}{6}}(2^{\frac{1}{2}}+2) \\ &\le x(5+2^{\frac{1}{2}}) + 2Q^2x^{\frac{1}{2}} + 2Q^{\frac{3}{2}}x^{\frac{2}{3}} + Qx^{\frac{5}{6}}(2^{\frac{1}{2}}+2),\end{aligned}$$

where we used $Q^2x^{\frac{1}{3}} \le x$, $Q^{\frac{5}{2}}x^{\frac{1}{3}} \le Q^{\frac{3}{2}}x^{\frac{2}{3}}$ and $Q^{\frac{1}{2}}x^{\frac{5}{6}} \le x$.

Similarly we get for

$$\mathrm{G}_2(x,V,U) \le 2\left(\frac{7}{6}\right)^{\frac{3}{2}}(\log x)^{\frac{5}{2}}.$$

Finally, we have in (28)

$$S \le c_4\frac{2^4}{3^{\frac{3}{2}}}(7x + 2Q^2x^{\frac{1}{2}} + 5Q^{\frac{3}{2}}x^{\frac{2}{3}} + 4Qx^{\frac{5}{6}})(\log x)^{\frac{5}{2}},$$

as demanded.

4.5 Finishing the proof of Theorem 4

Let $y \ge 2$,$(a,q) = 1$. By orthogonality of characters modulo q, we have

$$\psi(y;q,a) = \frac{1}{\varphi(q)}\sum_{\chi}\overline{\chi}(a)\psi(y,\chi).$$

Define $\psi'(y,\chi) = \psi(y,\chi)$ if $\chi \ne \chi_0$ and $\psi'(y,\chi) = \psi(y,\chi) - \psi(y)$ otherwise, χ_0 is the principal character mod q. Then

$$\psi(y,q,a) - \frac{\psi(y)}{\varphi(q)} = \frac{1}{\varphi(q)}\sum_{\chi}\overline{\chi}(a)\psi'(y,\chi).$$

For a character $\chi \pmod q$, we let χ^* be the primitive character modulo q^* inducing χ. Follow the way of [1] we obtain

$$\psi'(y,\chi^*) - \psi'(y,\chi) = \psi(y,\chi^*) - \psi(y,\chi) = \sum_{p^k \le y}(\log p)(\chi^*(p^k) - \chi(p^k)).$$

If $p|q$ then $(p^k,q^*) = 1$, and hence $\chi^*(p^k) = \chi(p^k)$. If $p|q$ then $\chi(p^k) = 0$. Therefore

$$|\psi'(y,\chi^*) - \psi'(y,\chi)| \le \sum_{\substack{p^k \le y \\ p|q}}(\log p) \le (\log y)\sum_{p|q}1 \le (\log qy)^2.$$

Denote the quantity we want to estimate as

$$\mathcal{M} = \sum_{\substack{q\le Q\\ l(q)>Q_1}} \max_{2\le y\le x} \max_{\substack{a\\(a,q)=1}} \left|\psi(y;q,a) - \frac{\psi(y)}{\varphi(q)}\right|.$$

Since

$$\left|\psi(y,q,a) - \frac{\psi(y)}{\varphi(q)}\right| \le \frac{1}{\varphi(q)}\sum_{\chi}|\psi'(y,\chi)| \le (\log qy)^2 + \frac{1}{\varphi(q)}\sum_{\chi}|\psi'(y,\chi^*)|,$$

then

$$\mathcal{M} \le Q(\log Qx)^2 + \sum_{\substack{q\le Q\\ l(q)>Q_1}} \frac{1}{\varphi(q)}\sum_{\chi}\max_{2\le y\le x}|\psi'(y,\chi^*)|.$$

We have to take care just of the second term in the inequality above, since the first one is smaller than the desired bound. It remains to prove

$$\mathcal{N} = \sum_{\substack{q\le Q\\ l(q)>Q_1}} \frac{1}{\varphi(q)}\sum_{\chi}\max_{2\le y\le x}|\psi'(y,\chi^*)| \le (c_1-1)F(x,Q,Q_1)(\log x)^4,$$

where $F(x,Q,Q_1)$ is the function from Theorem 4. A primitive character $\chi^* \bmod q^*$ induces characters of moduli dq^* and $\psi'(y,\chi^*) = 0$ for χ principal, we observe

$$\mathcal{N} = \sum_{\substack{q\le Q\\ l(q)>Q_1}} \frac{1}{\varphi(q)}\sum_{\substack{q^*|q\\ q^*\ne 1}}\sideset{}{^*}\sum_{\chi(q^*)}\max_{2\le y\le x}|\psi'(y,\chi)| \le \sum_{\substack{q^*\le Q\\ l(q^*)>Q_1}}\sideset{}{^*}\sum_{\chi(q^*)}\max_{2\le y\le x}|\psi'(y,\chi)|\sum_{k\le \frac{Q}{q^*}}\frac{1}{\varphi(kq^*)}.$$

As it was noted in [1] for $x > 0$

$$\sum_{k\le x}\frac{1}{\varphi(k)} \le E_0\log(ex)$$

and as $q^* \le Q \le x^{1/2}$, $\varphi(k)\varphi(q^*) \le \varphi(kq^*)$ and $x \ge 4$, we have

$$\sum_{k\le \frac{Q}{q^*}}\frac{1}{\varphi(kq^*)} < \frac{5E_0}{4\varphi(q^*)}\log x.$$

For $q > 1$ and χ primitive character $(\bmod q)$, we know that χ is non-principal and $\psi(y,\chi) = \psi'(y,\chi)$. Since we assumed $Q_1 \ge 1$ then we can can replace $\psi'(y,\chi)$ by $\psi(y,\chi)$ inside the internal sum for $\mathcal{N}$. Combining it with an expression for $\mathcal{N}$ we get

$$\mathcal{N} \le \frac{5E_0}{4}(\log x)\sum_{\substack{q\le Q\\ l(q)>Q_1}}\frac{1}{\varphi(q)}\sideset{}{^*}\sum_{\chi(q)}\max_{2\le y\le x}|\psi(y,\chi)| = \mathcal{R}.$$

Thus it remains to show that

$$\mathcal{R} \le \frac{4(c_1-1)}{5E_0}F(x,Q,Q_1)(\log x)^{\frac{5}{2}}.$$

Let

$$\mathcal{R}(q) = \frac{q}{\varphi(q)} \sum_{\chi(q)}^{*} \max_{2\le y\le x} |\psi(y,\chi)|.$$

Partial summation gives us

$$\sum_{Q_1<q\le Q} \frac{1}{\varphi(q)} \sum_{\chi(q)}^{*} \max_{2\le y\le x} |\psi(y,\chi)| = \frac{1}{Q}\sum_{q\le Q}\mathcal{R}(q) - \frac{1}{Q_1}\sum_{q\le Q_1}\mathcal{R}(q) + \int_{Q_1}^{Q}\left(\sum_{q\le t}\mathcal{R}(q)\right)\frac{dt}{t^2}.$$

Now we apply Theorem 1

$$\sum_{q\le Q}\mathcal{R}(q) < c_0 f(x,Q)(\log x)^{\frac{5}{2}},$$

where $f(x,Q) = 7x + 2Q^2x^{\frac{1}{2}} + 5Q^{\frac{3}{2}}x^{\frac{2}{3}} + 4Qx^{\frac{5}{6}}$. Then

$$\sum_{Q_1<q\le Q} \frac{1}{\varphi(q)} \sum_{\chi(q)}^{*} \max_{2\le y\le x} |\psi(y,\chi)| < c_0\left(\Delta_f(Q,Q_1) + \int_{Q_1}^{Q} f(x,t)\frac{dt}{t^2}\right)(\log x)^{\frac{5}{2}},$$

where

$$\Delta_f(Q,Q_1) = \frac{f(x,Q)}{Q} - \frac{f(x,Q_1)}{Q_1} \le \frac{7x}{Q_1} + 2x^{\frac{1}{2}}Q + 5x^{\frac{2}{3}}Q^{\frac{1}{2}}.$$

Calculating the integrals gives us

$$\int_{Q_1}^{Q} f(x,t)\frac{dt}{t^2} < \frac{7x}{Q_1} + 2x^{\frac{1}{2}}Q + 10x^{\frac{2}{3}}Q^{\frac{1}{2}} + 4x^{\frac{5}{6}}\log\frac{Q}{Q_1}.$$

Finally

$$\mathcal{N} \le \frac{4(c_1-1)}{5E_0}\left(\frac{14x}{Q_1} + 4x^{\frac{1}{2}}Q + 15x^{\frac{2}{3}}Q^{\frac{1}{2}} + 4x^{\frac{5}{6}}\log\frac{Q}{Q_1}\right)(\log x)^{\frac{5}{2}}.$$

4.6 Proof of Remark 1

Define two functions

$$\pi_1(y) = \sum_{2\le n\le y}\frac{\Lambda(n)}{\log n} \qquad \text{and} \qquad \pi_1(y;q,a) = \sum_{\substack{2\le n\le y\\ n\equiv a(\mathrm{mod}\, q)}}\frac{\Lambda(n)}{\log n}.$$

Since

$$\pi_1(y;q,a) - \pi(y;q,a) = \sum_{2\le k\le \frac{\log y}{\log 2}}\sum_{\substack{p^k\le y\\ p^k\equiv a(\mathrm{mod}\, q)}}\frac{1}{k} \le \sum_{2\le k\le \frac{\log y}{\log 2}}\frac{\pi(y^{\frac{1}{2}})}{2} < 2y^{\frac{1}{2}},$$

where we used the fact that for $x > 1$(see for example [1, Lemma 3.1])

$$\pi(x) < 1.25506\frac{x}{\log x}.$$

Similarly, $\pi_1(y) - \pi(y) < 2y^{\frac{1}{2}}$. Thus by partial summation we obtain the bound

$$\left|\pi_1(y;q,a) - \frac{\pi_1(y)}{\varphi(q)}\right| = \left|\frac{\psi(y;q,a) - \psi(y)/\varphi(q)}{\log y} - \int_2^y \frac{\psi(t;q,a) - \psi(t)/\varphi(q)}{t\log^2 t}dt\right|$$
$$\leq \frac{1}{\log 2}\left|\psi(y;q,a) - \frac{\psi(y)}{\varphi(q)}\right| + \max_{2\leq t\leq y}\left|\psi(t;q,a) - \frac{\psi(t)}{\varphi(q)}\right|\left(\frac{1}{\log 2} - \frac{1}{\log y}\right).$$

We have

$$\sum_{\substack{q\leq Q\\ l(q)>Q_1}} \max_{2\leq y\leq x}\max_{\substack{a\\(a,q)=1}}\left|\pi(y;q,a) - \frac{\pi(y)}{\varphi(q)}\right|$$
$$\leq \frac{2}{\log 2}\sum_{\substack{q\leq Q\\ l(q)>Q_1}} \max_{2\leq y\leq x}\max_{a,(a,q)=1}\left|\psi(y,q,a) - \frac{\psi(y)}{\varphi(q)}\right| + 2x^{\frac{1}{2}}\sum_{\substack{q\leq Q\\ l(q)>Q_1}}\left(1 + \frac{1}{\varphi(q)}\right)$$
$$\ll F(x,Q,Q_1)(\log x)^2 + x^{\frac{1}{2}}\sum_{\substack{q\leq Q\\ l(q)>Q_1}}\left(1 + \frac{1}{\varphi(q)}\right),$$

where we used Theorem 5 to estimate the first summand and denote

$$F(x,Q,Q_1) = x^{\frac{1}{2}}Q + \frac{x}{Q_1}(\log x) + x^{\frac{95}{96}}(\log x)^2.$$

For $x \geq 4$

$$x^{\frac{1}{2}}\sum_{\substack{q\leq Q\\ l(q)>Q_1}}\left(1 + \frac{1}{\varphi(q)}\right) \ll F(x,Q,Q_1)(\log x)^2.$$

and we are done.

5 Bombieri-Vinogradov theorem, an effective version

5.1 Auxiliary statements

In this section we establish an improved version of what we had before, namely, Theorem 4, by different techniques. In doing that we need a weighted version of Vaughan identity, see Lemma 4, that allows us to get further cancellation in one of type II sums via a method of Graham [17], see Lemmas 6 and 3. Next we use the large sieve inequality to get the cancellation in the remaining type II sums. We work with type I sums by standard means.

Our goal here is to prove a better version of Proposition 1, namely, Proposition 2 that is the key for getting Theorem 5.

Proposition 2 (Vaughan's inequality, improved)**.** *For $x \geq 4$ and any positive $\varepsilon < \frac{1}{28}$ we have*

$$\sum_{q\leq Q} \frac{q}{\varphi(q)} \sideset{}{^*}\sum_{\chi(q)} \max_{y\leq x} |\psi(y,\chi)| \ll \left(x + Q^{\frac{3}{2}} x^{\frac{9}{16}} + Q^2 x^{\frac{1}{2}} + Q x^{\frac{13}{14}+\varepsilon}\right) (\log x)^2,$$

where Q is any positive real number and $\sum^*_{\chi(q)}$ means a sum over all primitive characters $\chi (\mathrm{mod}\, q)$.

We formulate a weighted version of Vaughan identity.

Lemma 4 (Weighted Vaughan identity)**.** *Let $U, V \geq 1$, $n > U$. Define a function $\eta(t) : \mathbb{Z}^+ \to \mathbb{R}$ with $\eta(t) = 1$ for $t \leq V$. We have*

$$\Lambda(n) = \lambda_0(n) + \lambda_1(n) + \lambda_2(n) + \lambda_3(n),$$

where

$$\begin{aligned}
\lambda_0(n) &= \begin{cases} \Lambda(n), & n \leq U, \\ 0, & n > U, \end{cases} \\
\lambda_1(n) &= \sum_{d|n} \mu(d)\eta(d) \log \frac{n}{d}, \\
\lambda_2(n) &= -\sum_{c\leq U} \sum_{dc|n} \mu(d)\Lambda(c)\eta(d), \\
\lambda_3(n) &= \sum_{c>U} \sum_{dc|n} \mu(d)\Lambda(c)(1-\eta(d)).
\end{aligned}$$

Proof. Let $n > U$, since otherwise the statement is trivial. Define the following quantities

$$\begin{aligned}
\Lambda_1(n) &= \sum_{\substack{d|n \\ d\leq V}} \mu(d) \log \frac{n}{d} = \lambda_1(n) - \sum_{\substack{d|n \\ d>V}} \mu(d)\eta(d) \log \frac{n}{d} = \lambda_1(n) + \lambda_1'(n), \\
\Lambda_2(n) &= -\sum_{c\leq U} \sum_{\substack{dc|n \\ d\leq V}} \mu(d)\Lambda(c) = \lambda_2(n) + \sum_{c\leq U} \sum_{\substack{dc|n \\ d>V}} \mu(d)\Lambda(c)\eta(d) = \lambda_2(n) + \lambda_2'(n), \\
\Lambda_3(n) &= \sum_{c>U} \sum_{\substack{dc|n \\ d>V}} \mu(d)\Lambda(c) = \lambda_3(n) + \sum_{c>U} \sum_{\substack{dc|n \\ d>V}} \mu(d)\Lambda(c)\eta(d) = \lambda_3(n) + \lambda_3'(n).
\end{aligned}$$

Vaughan's identity in its classical form is

$$\Lambda(n) = \Lambda_1(n) + \Lambda_2(n) + \Lambda_3(n),$$

so it remains to show that $\lambda_1'(n) + \lambda_2'(n) + \lambda_3'(n) = 0$ for every n. Let us rewrite this sum

$$\begin{aligned}\sum_{i=1}^{3} \lambda_i'(n) &= \sum_{\substack{d|n \\ d>V}} \left(-\mu(d)\eta(d)\log\frac{n}{d} + \sum_{\substack{c|\frac{n}{d} \\ c\le U}} \mu(d)\Lambda(c)\eta(d) + \sum_{\substack{c|\frac{n}{d} \\ c>U}} \mu(d)\Lambda(c)\eta(d) \right) \\ &= \sum_{\substack{d|n \\ d>V}} \left(-\mu(d)\eta(d)\log\frac{n}{d} + \mu(d)\eta(d)\sum_{c|\frac{n}{d}} \Lambda(c) \right) = 0,\end{aligned}$$

where in the last equality we used the fact that $\sum_{x|y} \Lambda(x) = \log y$. □

Lemma 6 (Graham [17])**.** *Let $1 \le N_1 \le N_2 \le N$ and define $f_i(d) = \mu(d)\log\frac{N_i}{d}$ for $d \le N_i$ and $f_i(d) = 0$ for $d > N_i$. Then we have*

$$\sum_{n=1}^{N} \left(\sum_{d_1|n} f_1(d_1) \right) \left(\sum_{d_2|n} f_2(d_2) \right) = N\log N_1 + O(N).$$

For the proof see the paper of Graham [17]. From the lemma above one can deduce

Corollary 3. *Define a function $\eta(t)$, that is equal to 1 for $t \le V$, to 0 for $t > V_0$ and*

$$\eta(t) = \frac{\log\frac{V_0}{t}}{\log\frac{V_0}{V}}, \quad V < t \le V_0.$$

Then

$$\sum_{k\le Y} \left| \sum_{d|k} \mu(d)\eta(d) \right|^2 \ll \frac{Y}{\log\frac{V_0}{V}}.$$

The constant here can be made explicit as in [22].

5.2 Proof of Proposition 2

We proceed now with the proof of Proposition 1. Fix arbitrary real numbers $Q > 0$ and $x \ge 4$.

Without loss of generality we can assume that $2 \le Q \le x^{1/2}$ and decompose the von Mangoldt function using a weighted form of Vaughan's identity, namely Lemma 4.

$$\Lambda(n) = \lambda_0(n) + \lambda_1(n) + \lambda_2(n) + \lambda_3(n),$$

where $\lambda_i(n)$, $i = 0, 1, 2, 3$ are as in the statement of the lemma and U, V, V_0 are parameters. Notice also that we are free to choose $\eta(t)$ as we wish, we only need to fulfill the conditions stated in Lemma 4.

Assume $y \le x$, $q \le Q$, and χ is a character mod q. We use the above decomposition to write

$$\psi(y,\chi) = s_0 + s_1 + s_2 + s_3,$$

where

$$s_i = \sum_{n\le y} \lambda_i(n)\chi(n).$$

Denote the contributions to our main sum by

$$S_i = \sum_{q\le Q} \frac{q}{\varphi(q)} \sideset{}{^*}\sum_{\chi(q)} \max_{y\le x} |s_i|.$$

Easily we obtain

$$\sum_{q\le Q} \frac{q}{\varphi(q)} \sideset{}{^*}\sum_{\chi(q)} \max_{y\le x} |\psi(y,\chi)| \le S_0 + S_1 + S_2 + S_3,$$

where

$$\begin{aligned}
&S_0 \ll UQ^2,\\
&s_1 = \sum_{d\le y} \mu(d)\chi(d)\eta(d) \sum_{h\le \frac{y}{d}} \chi(h)\log h,\\
&s_2 = -\sum_{\substack{dcr\le y\\ c\le U}} \chi(dcr)\mu(d)\Lambda(c)\eta(d),\\
&s_3 = \sum_{n\le y} \chi(n) \sum_{c>U} \sum_{dc|n} \mu(d)\Lambda(c)(1-\eta(d)).
\end{aligned} \tag{29}$$

Here in bounding S_0 we used Chebychev's estimate

$$|s_0| \le \sum_{n\le U} \Lambda(n) \ll U.$$

In what follows we choose $\eta(\cdot)$ from the paper by Graham, see [17]:

$$\eta(d) = \frac{\log \frac{V_0}{d}}{\log \frac{V_0}{V}}, \quad V \le d \le V_0.$$

We remind that $\eta(d) = 1$ for $d \le V$ and $\eta(d) = 0$ for $d > V_0$. This choice allows us to win $\log^{\frac{1}{2}}$ in the last sum, that is of type II.

5.2.1 Type I sums

We start with linear sums among s_i and work with s_1 first. Write

$$\sum_{h\le \frac{y}{d}} \chi(h)\log h = \sum_{h\le \frac{y}{d}} \chi(h) \int_1^h \frac{du}{u}$$

and exchange the sum and the integral

$$\begin{aligned}
s_1 &= \sum_{d\le V_0}\mu(d)\chi(d)\eta(d)\int_1^{\frac{y}{d}}\sum_{u\le h\le\frac{y}{d}}\chi(h)\frac{du}{u}\\
&= \int_1^{y}\sum_{d\le\min(V_0,\frac{y}{u})}\mu(d)\chi(d)\eta(d)\sum_{u\le h\le\frac{y}{d}}\chi(h)\frac{du}{u}\\
&= \int_1^{y}\left(\sum_{d\le V_0}\mu(d)\chi(d)\eta(d)\sum_{u\le h\le\frac{y}{d}}\chi(h)\right)\frac{du}{u}\\
&= \int_1^{y}\left(\sum_{d\le V}\mu(d)\chi(d)\sum_{u\le h\le\frac{y}{d}}\chi(h)\right)\frac{du}{u}\\
&+ \frac{1}{\log\frac{V_0}{V}}\int_1^{y}\left(\sum_{d\le V_0}\mu(d)\chi(d)\log\frac{V_0}{d}\sum_{u\le h\le\frac{y}{d}}\chi(h)\right)\frac{du}{u}.
\end{aligned}$$

Denote the summands σ_1 and σ_2. Then

$$|\sigma_1|\le\sum_{d\le V}\max_{1\le u\le y}\left|\sum_{u\le h\le\frac{y}{d}}\chi(h)\right|\int_1^{y}\frac{du}{u}\le(\log y)\sum_{d\le V}\max_{1\le u\le y}\left|\sum_{u\le h\le\frac{y}{d}}\chi(h)\right|.$$

If $q=1$, then we have only trivial $\chi \bmod q$ and

$$|\sigma_1|\le(\log y)\sum_{d\le V}\frac{1}{d}\le x(\log xV)^2.$$

If $q>1$ and χ is a primitive character $\bmod q$, we use the Pólya-Vinogradov inequality: for all x,y we have

$$\left|\sum_{x\le n\le y}\chi(n)\right|<q^{\frac12}\log q.$$

Then

$$|\sigma_1|<(\log y)\sum_{d\le V}\max_{1\le u\le y}q^{\frac12}\log q<q^{\frac12}V(\log xV)^2.$$

Further

$$|\sigma_2|\le\frac{\log V_0}{\log\frac{V_0}{V}}|\sigma_1|$$

and

$$\begin{aligned}
S_1 &\le \left(\log\frac{V_0}{V}\right)^{-1}\log\frac{V_0^2}{V}\left(\sum_{\chi \bmod q=1}^{*}\max_{y\le x}|s_1|+\sum_{1<q\le Q}\frac{q}{\varphi(q)}\sum_{\chi(q)}^{*}\max_{y\le x}|s_1|\right)\\
&\le \left(\log\frac{V_0}{V}\right)^{-1}\log\frac{V_0^2}{V}\left(x(\log xV)^2+V(\log xV)^2\sum_{1<q\le Q}\frac{q^{\frac32}}{\varphi(q)}\sum_{\chi(q)}^{*}1\right)\\
&\le (x+Q^{\frac52}V)(\log xV)^2\left(\log\frac{V_0}{V}\right)^{-1}\log\frac{V_0^2}{V}.
\end{aligned}$$

5.2.2 Type II sums

Now we work with s_2 and want to use dyadic decomposition. Write

$$s_2 = -\sum_{\substack{cdr\le y\\ c\le U}} \Lambda(c)\mu(d)\eta(d)\chi(cdr) = -\sum_{\substack{ct\le y\\ c\le U}}\sum_{d|t} \Lambda(c)\mu(d)\eta(d)\chi(ct)$$
$$= -\sum_{c\le w} - \sum_{w<c\le U} = s_2' + s_2'',$$

where we introduced a new parameter w, that should be smaller than U and will be chosen later. We deal first with the linear part of s_2, namely s_2'. Write

$$s_2' = -\sum_{c\le w}\Lambda(c)\chi(c)\sum_{t\le \frac{y}{c}}\sum_{d|t}\mu(d)\eta(d)\chi(t).$$

Since we have the bound

$$\left|\sum_{\substack{cd=t\\ c\le w}}\Lambda(c)\mu(d)\eta(d)\chi(t)\right| \le \sum_{c|t}\Lambda(c) = \log t,$$

then proceeding as for s_1 via Pólya-Vinogradov inequality and using the fact that $cd = t \le wV_0$ we get

$$|S_2'| \le (x + Q^{\frac{5}{2}}wV_0)(\log(xwV_0))^2,$$

where the x term comes from the contribution of $q = 1$ and $Q^{\frac{5}{2}}wV_0$ from the remaining $q \ne 1$.

Next consider s_2''. Writing s_2'' as a dyadic sum we have

$$s_2'' = \sum_{\substack{M=2^\alpha\\ \frac{1}{2}w<M\le U}}\sum_{\substack{w<c\le U\\ M<c\le 2M}}\sum_{t\le\frac{y}{c}}\sum_{d|t}\Lambda(c)\mu(d)\eta(d)\chi(ct).$$

Using the triangle inequality

$$S_2'' \le \sum_{\substack{M=2^\alpha\\ \frac{1}{2}w<M\le U}}\sum_{q\le Q}\frac{q}{\varphi(q)}\sideset{}{^*}\sum_{\chi(q)}\max_{y\le x}\left|\sum_{\substack{w<c\le U\\ M<c\le 2M}}\sum_{t\le\frac{y}{c}}\sum_{d|t}\Lambda(c)\mu(d)\eta(d)\chi(ct)\right|.$$

By the large sieve inequality we get

$$S_2'' \ll \sum_{\substack{M=2^\alpha\\ \frac{1}{2}w<M\le U}}(M'+Q^2)^{\frac{1}{2}}(K'+Q^2)^{\frac{1}{2}}\sigma_1(M)^{\frac{1}{2}}\sigma_2(M)^{\frac{1}{2}}L(M),$$

where M' and K' are the number of terms in sums over c and t respectively and

$$\sigma_1(M) = \sum_{\substack{w<c<U\\ M<c\le 2M}}\Lambda(c)^2,$$

$$\sigma_2(M) = \sum_{t\le\frac{y}{M}}\left|\sum_{d|t}\mu(d)\eta(d)\right|^2,$$

and

$$L(M) = \log\left(\frac{2x}{M}\min(U, 2M)\right) \ll \log x,$$

By Chebyshev's estimate

$$\sigma_1(M) \ll 2M\log U,$$

then using the estimates $M' \le M$, $K' \le \frac{x}{M}$ we have

$$S_2'' \ll (\log x)(\log U)^{\frac{1}{2}} \sum_{\substack{M=2^\alpha \\ \frac{1}{2}w<M\le U}} (M+Q^2)^{\frac{1}{2}}\left(\frac{x}{M}+Q^2\right)^{\frac{1}{2}} M^{\frac{1}{2}}\sigma_2(M)^{\frac{1}{2}}.$$

To bound $\sigma_2(M)$ we use a result of Corollary 3 and get

$$\sigma_2(M) \ll \frac{y}{M\log\frac{V_0}{V}}.$$

Putting it together we obtain

$$\begin{aligned} S_2'' &\ll (\log x)(\log U)^{\frac{1}{2}} x^{\frac{1}{2}} \sum_{\substack{M=2^\alpha \\ \frac{1}{2}w<M\le U}} (M+Q^2)^{\frac{1}{2}}\left(\frac{x}{M}+Q^2\right)^{\frac{1}{2}}\left(\log\frac{V_0}{V}\right)^{-\frac{1}{2}} \\ &\ll (\log x)\frac{(\log U)^{\frac{1}{2}}}{\left(\log\frac{V_0}{V}\right)^{\frac{1}{2}}}(\log(Uw))\left(x+\sqrt{2}Qxw^{-\frac{1}{2}}+Q^2x^{\frac{1}{2}}+U^{\frac{1}{2}}Qx^{\frac{1}{2}}\right), \end{aligned}$$

where we applied the bound

$$\sum_{\substack{M=2^\alpha \\ \frac{1}{2}w<M\le U}} 1 \le \frac{\log(2Uw)}{\log 2}.$$

We continue with an estimate for S_3 and use of the large sieve inequality (11) and properties of $\eta(\cdot)$ from Lemma 6. Writing s_3 as a dyadic sum we have

$$s_3 = -\sum_{\substack{M=2^\alpha \\ \frac{1}{2}U<M\le x/V}} \sum_{\substack{U<m\le x/V \\ M<m\le 2M}} \sum_{\substack{V<k\le x/M \\ mk\le y}} \Lambda(m)\left(\sum_{\substack{d\mid k \\ d\le V}} \mu(d)(1-\eta(d))\right)\chi(mk).$$

Using the triangle inequality

$$S_3 \le \sum_{\substack{M=2^\alpha \\ \frac{1}{2}U<M\le x/V}} \sum_{q\le Q} \frac{q}{\varphi(q)} \sideset{}{^*}\sum_{\chi(q)} \max_{y\le x}\left|\sum_{\substack{U<m\le x/V \\ M<m\le 2M}} \sum_{\substack{V<k\le x/M \\ mk\le y}} a_m c_k \chi(mk)\right|,$$

where $a_m = \Lambda(m)$ and $c_k = \sum_{d\mid k,\, d\le V}\mu(d)(1-\eta(d))$. Now apply the large sieve inequality (11) to get

$$S_3 \ll \sum_{\substack{M=2^\alpha \\ \frac{1}{2}U<M\le x/V}} (M'+Q^2)^{\frac{1}{2}}(K'+Q^2)^{\frac{1}{2}}\sigma_1(M)^{\frac{1}{2}}\sigma_2(M)^{\frac{1}{2}}L(M)$$

where

$$\sigma_1(M) = \sum_{V<k\le x/M} |c_k|^2,$$

$$\sigma_2(M) = \sum_{\substack{U<m\le x/V \\ M<m\le 2M}} |a_m|^2,$$

and

$$L(M) = \log\left(\frac{2x}{M}\min\left(\frac{x}{V}, 2M\right)\right) \ll \log x,$$

where M' and K' denote the number of terms in the sums over m and k, respectively. From the definition of M' and N' we conclude

$$M' = \min\left(2M, \frac{x}{V}\right) - \max\left(M+1, U+1\right) \le M,$$
$$K' = \frac{x}{M} - (V+1) + 1 \le \frac{x}{M}.$$

By Chebyshev's estimate we have an upper bound

$$\sigma_2(M) \le \sum_{m\le 2M} \Lambda(m)^2 \le \psi(2M)\log 2M \ll M\log M.$$

Thus by Cauchy inequality

$$S_3 \ll (\log x) \sum_{\substack{M=2^\alpha \\ \frac{1}{2}U<M\le x/V}} (M+Q^2)^{\frac{1}{2}} \left(\frac{x}{M}+Q^2\right)^{\frac{1}{2}} (M\log M)^{\frac{1}{2}}\sigma_1(M)^{\frac{1}{2}}.$$

Further

$$M(M+Q^2)\left(\frac{x}{M}+Q^2\right) = Mx + Q^2x + M^2Q^2 + MQ^4$$

and

$$(\log M)^{\frac{1}{2}} \le \left(\log\frac{x}{V}\right)^{\frac{1}{2}}.$$

Thus we have

$$S_3 \ll (\log x)\left(\log\frac{x}{V}\right)^{\frac{1}{2}} \sum_{\substack{M=2^\alpha \\ \frac{1}{2}U<M\le x/V}} (Mx + Q^2x + M^2Q^2 + MQ^4)^{\frac{1}{2}}\sigma_1(M)^{\frac{1}{2}}.$$

We take $\eta(\cdot)$ from the paper by Graham, see Corollary 3 and [17]:

$$\eta(d) = \begin{cases} 1, & d \le V, \\ \frac{\log V_0/d}{\log V_0/V}, & V \le d \le V_0, \\ 0, & d \ge V_0. \end{cases}$$

so that

$$1 - \eta(d) = 1 - \frac{\log\frac{V_0}{d}}{\log\frac{V_0}{V}} = \frac{\log\frac{d}{V}}{\log\frac{V_0}{V}}.$$

On applying Lemma 6 we obtain

$$\sigma_1(M) = \sum_{V<k\le \frac{x}{M}} \left(\sum_{d|k} \mu(d) - \sum_{d|k} \mu(d)\eta(d) \right)^2 \ll \frac{\log V}{\left(\log \frac{V_0}{V}\right)^2} \frac{x}{M}$$

that implies

$$S_3 \ll (\log x) \left(\log \frac{x}{V}\right)^{\frac{1}{2}} \frac{(\log V)^{\frac{1}{2}}}{\log \frac{V_0}{V}} \sum_{\substack{M=2^\alpha \\ \frac{1}{2}U<M\le \frac{x}{V}}} \left(x^2 + \frac{Q^2x^2}{M} + MQ^2x + Q^4x \right)^{\frac{1}{2}}.$$

Since

$$\sum_{\substack{M=2^\alpha \\ \frac{1}{2}U<M\le x/V}} 1 \le \frac{\log \frac{2x}{V}}{\log 2},$$

then

$$S_3 \ll \frac{\log x}{\log \frac{V_0}{V}} \left(\log \frac{x}{V}\right)^{\frac{3}{2}} (\log V)^{\frac{1}{2}} \left(x + \sqrt{2}QxU^{-\frac{1}{2}} + QxV^{-\frac{1}{2}} + Q^2x^{\frac{1}{2}} \right).$$

Finally we have to adjust the parameters U, V, V_0, w. We repeat our previous estimates

$$\begin{aligned}
&S_0 \ll UQ^2,\\
&S_1 \le (x + Q^{\frac{5}{2}}V)(\log xV)^2 \left(\log \frac{V_0}{V}\right)^{-1} \log \frac{V_0^2}{V},\\
&S_2' \le (x + Q^{\frac{5}{2}}wV_0)(\log(xwV_0))^2,\\
&S_2'' \ll (\log x)\frac{(\log U)^{\frac{1}{2}}}{\left(\log \frac{V_0}{V}\right)^{\frac{1}{2}}}(\log(2Uw))(x + \sqrt{2}Qxw^{-\frac{1}{2}} + Q^2x^{\frac{1}{2}} + U^{\frac{1}{2}}Qx^{\frac{1}{2}}),\\
&S_3 \ll \frac{\log x}{\log \frac{V_0}{V}} \left(\log \frac{x}{V}\right)^{\frac{3}{2}} (\log V)^{\frac{1}{2}}(x + \sqrt{2}QxU^{-\frac{1}{2}} + QxV^{-\frac{1}{2}} + Q^2x^{\frac{1}{2}}).
\end{aligned}$$

Combining the results above and taking $U = V$ we get

$$S = \sum_{q\le Q} \frac{q}{\varphi(q)} \sideset{}{^*}\sum_{\chi(q)} \max_{y\le x} |\psi(y,\chi)| \ll \mathrm{R}(x,Q,w,V,V_0)\,\mathrm{G}(x,w,V,V_0), \tag{30}$$

where

$$\begin{aligned}
\mathrm{R}(x,Q,w,V,V_0) &= 4x + Q^2V + Q^{\frac{5}{2}}(V + wV_0) + Qx\left(\frac{\sqrt{2}}{w^{\frac{1}{2}}} + \frac{1+\sqrt{2}}{V^{\frac{1}{2}}}\right)\\
&\quad + 2Q^2x^{\frac{1}{2}} + V^{\frac{1}{2}}Qx^{\frac{1}{2}},\\
\mathrm{G}(x,Q,w,V,V_0) &= \max\{(\log xV)^2\frac{\log \frac{V_0^2}{V}}{\log \frac{V_0}{V}}, (\log(xwV_0))^2, (\log(Vw))\frac{(\log V)^{\frac{1}{2}}}{\left(\log \frac{V_0}{V}\right)^{\frac{1}{2}}}\log x,\\
&\quad \left(\log \frac{2x}{V}\right)^{\frac{3}{2}} \frac{\log 4x}{\log \frac{V_0}{V}}(\log V)^{\frac{1}{2}}\}.
\end{aligned}$$

Now let's specify V and V_0. We introduce a parameter $0 < \alpha < \frac{1}{2}$ to be chosen later. We subdivide into two cases

1. $x^{\alpha} \le Q \le x^{\frac{1}{2}}$,

2. $Q \le x^{\alpha}$

and denote $\mathrm{R}(x,Q,w,V,V_0)$, $\mathrm{G}(x,w,V,V_0)$ as $\mathrm{R}_1, \mathrm{R}_2$ and, respectively G_1 and G_2. If $x^{\alpha} \le Q \le x^{\frac{1}{2}}$, then $V = \frac{x^{\beta_1}}{Q}$. We choose $V_0 = \frac{x^{\delta_1}}{Q}$ and $w = \frac{x^{\gamma_1}}{Q}$. Then putting that into previous expression $\mathrm{R}(x,Q,w,V,V_0)$ we get for the factor

$$\begin{aligned}\mathrm{R}_1(x,Q) \ll x + Qx^{\beta_1} + Q^{\frac{3}{2}}x^{\beta_1} + Q^{\frac{1}{2}}x^{\gamma_1+\delta_1} + Q^{\frac{3}{2}}x^{1-\frac{\gamma_1}{2}} \\ + Q^{\frac{3}{2}}x^{1-\frac{\beta_1}{2}} + Q^2x^{\frac{1}{2}} + Q^{\frac{1}{2}}x^{\frac{1+\beta_1}{2}}.\end{aligned}$$

If $Q \le x^{\alpha}$, we let $V = x^{\beta_2}$, $V_0 = x^{\delta_2}$, $w = x^{\gamma_2}$ and get

$$\mathrm{R}_2(x,Q) \ll x + Q^2x^{\beta_2} + Q^{\frac{5}{2}}x^{\beta_2} + Q^{\frac{5}{2}}x^{\gamma_2+\delta_2} + Qx^{1-\frac{\gamma_2}{2}} + Qx^{1-\frac{\beta_2}{2}} + Q^2x^{\frac{1}{2}} + Qx^{\frac{1+\beta_2}{2}}.$$

Let $0 < \varepsilon < \frac{1}{14}$. We keep in mind conditions $\alpha < \frac{1}{2}$, $\gamma_1 < \beta_1$, $\delta_1 > \beta_1$ and put

$$\alpha = \frac{3}{7} + \varepsilon, \;\; \beta_1 = \frac{4}{7}, \;\; \gamma_1 = \frac{4}{7} - \varepsilon, \;\; \delta_1 = \frac{4}{7} + \frac{5\varepsilon}{2}.$$

Then

$$\begin{aligned}\mathrm{R}_1(x,Q) &\ll x + Qx^{\frac{4}{7}} + Q^{\frac{3}{2}}x^{\frac{4}{7}} + Q^{\frac{1}{2}}x^{\frac{8}{7}+\frac{3\varepsilon}{2}} + Q^{\frac{3}{2}}x^{\frac{5}{7}+\frac{\varepsilon}{2}} + Q^{\frac{3}{2}}x^{\frac{5}{7}} + Q^2x^{\frac{1}{2}} + Q^{\frac{1}{2}}x^{\frac{9}{14}} \\ &\ll x + Q^2x^{\frac{1}{2}},\end{aligned}$$

where we used

$$\begin{aligned}&Qx^{\frac{4}{7}} \le Q^2x^{\frac{4}{7}-\frac{3}{7}-\varepsilon} < Q^2x^{\frac{1}{2}}, \\ &Q^{\frac{3}{2}}x^{\frac{4}{7}} \le Q^2x^{\frac{4}{7}-\frac{3}{14}-\frac{\varepsilon}{2}} < Q^2x^{\frac{1}{2}}, \\ &Q^{\frac{1}{2}}x^{\frac{8}{7}+\frac{3\varepsilon}{2}} \le Q^2x^{\frac{8}{7}+\frac{3\varepsilon}{2}-\frac{3}{2}\left(\frac{3}{7}+\varepsilon\right)} = Q^2x^{\frac{1}{2}}, \\ &Q^{\frac{3}{2}}x^{\frac{5}{7}+\frac{\varepsilon}{2}} \le Q^2x^{\frac{5}{7}+\frac{\varepsilon}{2}-\frac{3}{14}-\frac{\varepsilon}{2}} = Q^2x^{\frac{1}{2}}, \\ &Q^{\frac{3}{2}}x^{\frac{5}{7}} \le Q^2x^{\frac{5}{7}-\frac{1}{2}\left(\frac{3}{7}+\varepsilon\right)} = Q^2x^{\frac{1}{2}-\frac{\varepsilon}{2}} < Q^2x^{\frac{1}{2}}, \\ &Q^{\frac{1}{2}}x^{\frac{9}{14}} \le Q^2x^{\frac{9}{14}-\frac{3}{2}\left(\frac{3}{7}+\varepsilon\right)} < Q^2x^{\frac{1}{2}}.\end{aligned}$$

Similarly to satisfy $\gamma_2 < \beta_2$, $\delta_2 > \beta_2$ we put

$$\beta_2 = \frac{1}{7}, \;\; \gamma_2 = \frac{1}{7} - \varepsilon, \;\; \delta_2 = \frac{1}{7} + \frac{\varepsilon}{2}$$

we obtain

$$\begin{aligned}\mathrm{R}_2(x,Q) &\ll x + Q^2x^{\frac{1}{7}} + Q^{\frac{5}{2}}x^{\frac{1}{7}} + Q^{\frac{5}{2}}x^{\frac{2}{7}-\frac{\varepsilon}{2}} + Qx^{\frac{13}{14}+\frac{\varepsilon}{2}} + Qx^{\frac{13}{14}} + Q^2x^{\frac{1}{2}} + Qx^{\frac{4}{7}} \\ &\ll x + Q^2x^{\frac{1}{2}} + Qx^{\frac{13}{14}+\frac{\varepsilon}{2}},\end{aligned}$$

where we used

$$\begin{aligned}&Q^{\frac{5}{2}}x^{\frac{1}{7}} \le Q^2x^{\frac{1}{7}+\frac{3}{14}+\frac{\varepsilon}{2}} = Q^2x^{\frac{5}{14}+\frac{\varepsilon}{2}} < Q^2x^{\frac{1}{2}}, \\ &Q^{\frac{5}{2}}x^{\frac{2}{7}-\frac{\varepsilon}{2}} \le Q^2x^{\frac{2}{7}-\frac{\varepsilon}{2}+\frac{3}{14}+\frac{\varepsilon}{2}} = Q^2x^{\frac{1}{2}}.\end{aligned}$$

Now we bound $\mathrm{G}(x,Q,w,V,V_0)$. We notice that with our choice of parameters above $\log\frac{V_0}{V} \gg \log x$, where the implied constant depends on β_i, δ_i. Thus $\mathrm{G}_1 \ll (\log x)^2$ and similarly $\mathrm{G}_2 \ll (\log x)^2$. Finally, we have

$$\sum_{q\le Q}\frac{q}{\varphi(q)}\sideset{}{^*}\sum_{\chi(q)}\max_{y\le x}|\psi(y,\chi)| \ll (x+Q^2x^{\frac12}+Qx^{\frac{13}{14}+\frac{\varepsilon}{2}})(\log x)^2.$$

The power $\frac{13}{14}+\frac{\varepsilon}{2}$ is optimal here. Indeed, let us show first that $\alpha > \frac37$. The system

$$\begin{cases} Q^{\frac12}x^{\gamma_1+\delta_1} \le Q^2x^{\frac12}, \\ Q^{\frac32}x^{1-\frac{\gamma_1}{2}} \le Q^2x^{\frac12}, \end{cases}$$

brings us to

$$\begin{cases} \gamma_1+\delta_1-\frac{3\alpha}{2} \le \frac12, \\ 1-\frac{\gamma_1}{2}-\frac{\alpha}{2} \le \frac12. \end{cases}$$

Solving this we obtain $\delta_1 \le \frac{5\alpha}{2}-\frac12$. Further since $Q^{\frac32}x^{1-\frac{\beta_1}{2}} \le Q^2x^{\frac12}$, we get

$$1-\alpha \le \beta_1 < \delta_1 \le \frac{5\alpha}{2}-\frac12.$$

Thus $\alpha > \frac37$. We use that to obtain the fact that the term Qx^A has $A > \frac{13}{14}$. Since $Q^{\frac52}x^{\gamma_2+\delta_2} \le Q^2x^{\frac12}$, we get $\gamma_2+\delta_2 \le \frac12-\frac{\alpha}{2} < \frac27$. The inequality $Qx^{1-\frac{\beta_2}{2}} \le Qx^A$ gives us $\delta_2 > \beta_2 \ge 2(1-A)$. Similarly for γ_2 we obtain $\gamma_2 \ge 2(1-A)$ because of the term $Qx^{1-\frac{\gamma_2}{2}} \le Qx^A$. Combining all of this we get

$$4(1-A) < \gamma_2+\delta_2 < \frac27$$

and thus $A > \frac{13}{14}$.

5.3 Finishing the proof of Theorem 5

In the section we prove Theorem 5 using Proposition 2. We use the standard notations

$$E(x;q,a) = \psi(x;q,a) - \frac{x}{\varphi(q)}$$

and

$$E(x,q) = \max_{(a,q)=1}|E(x;q,a)|, \quad E^*(x,q) = \max_{y\le x}E(y,q).$$

Invoking $(a,q)=1$ and using the orthogonality of characters one can write

$$\begin{aligned}\psi(y;q,a) &= \sum_{n\le y}\Lambda(n)1_{q,a}(n) = \sum_{n\le y}\Lambda(n)\frac{1}{\varphi(q)}\sum_{\chi \bmod q}\overline{\chi}(a)\chi(n) \\ &= \frac{1}{\varphi(q)}\sum_{\chi \bmod q}\overline{\chi}(a)\psi(y,\chi),\end{aligned}$$

where

$$1_{q,a}(n) = \begin{cases} 1, n \equiv a(\mathrm{mod}\, q), \\ 0, \text{ otherwise.} \end{cases}$$

Define $\psi'(y,\chi) = \psi(y,\chi) - 1_{\chi_0}(\chi)y$, where $1_{\chi_0}(\chi) = 1$ if $\chi = \chi_0$ and is 0 otherwise. Then we can write

$$\psi(y;q,a) - \frac{y}{\varphi(q)} = \frac{1}{\varphi(q)} \sum_{\chi \bmod q} \overline{\chi}(a)\psi'(y,\chi).$$

Applying triangle inequality gives

$$|E(y;q,a)| \leq \frac{1}{\varphi(q)} \sum_{\chi \bmod q} |\psi'(y,\chi)|$$

and since the right hand side does not depend on a we get

$$E(y,q) \leq \frac{1}{\varphi(q)} \sum_{\chi \bmod q} |\psi'(y,\chi)|.$$

Let χ be a character modulo q and χ_1 be the primitive character with a period q_1 that induces χ.

$$\psi'(y,\chi_1) - \psi'(y,\chi) = \psi(y,\chi_1) - \psi(y,\chi) = \sum_{\substack{p|q \\ p\nmid q_1}} \sum_{k \leq \frac{\log y}{\log p}} \chi_1(p^k)\log p$$

$$= O\left(\log y \sum_{p|q} \log p\right) = O((\log y)(\log q)) = O((\log qy)^2).$$

Then

$$E(y,q) \ll (\log qy)^2 + \frac{1}{\varphi(q)} \sum_{\chi \bmod q} |\psi'(y,\chi_1)|.$$

Using above we get that

$$\sum_{\substack{q\leq Q \\ l(q)>Q_1}} E^*(x,q) \ll \sum_{\substack{q\leq Q \\ l(q)>Q_1}} \left((\log qx)^2 + \frac{1}{\varphi(q)} \sum_{\chi \bmod q} \max_{y\leq x} |\psi'(y,\chi_1)|\right)$$

$$\leq Q(\log Qx)^2 + \sum_{\substack{q\leq Q \\ l(q)>Q_1}} \sum_{\chi \bmod q} \frac{1}{\varphi(q)} \max_{y\leq x} |\psi'(y,\chi_1)|. \tag{31}$$

A number that has all prime factors bigger than Q_1 is either 1 or is greater than or equal to Q_1. Thus we consider χ_0 separately and apply the prime number theorem to get

$$\sum_{\substack{q\leq Q \\ l(q)>Q_1}} E^*(x,q) \ll Q(\log Qx)^2 + \sum_{Q_1 < q_1 \leq Q} \sideset{}{^*}\sum_{\chi_1 \bmod q_1} \max_{y\leq x} |\psi'(y,\chi_1)| \left(\sum_{1\leq q'\leq Q/q_1} \frac{1}{\varphi(q'q_1)}\right)$$

$$+ \max_{y\leq x} |\psi'(y,\chi)| \sum_{q\leq Q} \frac{1}{\varphi(q)}$$

$$\ll Q(\log Qx)^2 + x(\log Q)e^{-c\sqrt{\log x}}$$

$$+ \sum_{Q_1 < q_1 \leq Q} \sideset{}{^*}\sum_{\chi_1 \bmod q_1} \max_{y\leq x} |\psi'(y,\chi_1)| \left(\sum_{1\leq q'\leq Q/q_1} \frac{1}{\varphi(q'q_1)}\right),$$

where the constant c is effective.

We follow the way of [47] and use the estimate

$$\sum_{\substack{m|q\\ m\le Q}} \frac{1}{\varphi(m)} \ll \frac{1}{\varphi(q)}\left(1+\log\frac{Q}{q}\right).$$

We have

$$\begin{aligned}
&\sum_{\substack{q\le Q\\ l(q)>Q_1}} E^*(x,q) \ll Q(\log Qx)^2 + x(\log Q)e^{-c\sqrt{\log x}}\\
&+\sum_{Q_1<q\le Q}\left(1+\log\frac{Q}{q}\right)\frac{1}{\varphi(q)}\sideset{}{^*}\sum_{\chi \bmod q}\max_{y\le x}|\psi'(y,\chi)|\\
&\ll Q(\log Qx)^2 + x(\log Q)e^{-c\sqrt{\log x}} + \frac{J(Q)}{Q} + \int_{Q_1}^{Q}\frac{J(\lambda)}{\lambda^2}\left(2+\log\frac{Q}{\lambda}\right)d\lambda,
\end{aligned}$$

where

$$J(\lambda) = \sum_{Q_1<q\le\lambda}\frac{q}{\varphi(q)}\sideset{}{^*}\sum_{\chi \bmod q}\max_{y\le x}|\psi'(y,\chi)|.$$

Using Proposition 2 for $J(\lambda)$ we get

$$\int_{Q_1}^{Q}\frac{J(\lambda)}{\lambda^2}\left(2+\log\frac{Q}{\lambda}\right)d\lambda = (\log x)^2\int_{Q_1}^{Q}\left(2+\log\frac{Q}{\lambda}\right)\left(\frac{x}{\lambda^2}+\frac{x^{\frac{13}{14}+\varepsilon}}{\lambda}\right)d\lambda.$$

Using the following calculations

$$\begin{aligned}
\int_{Q_1}^{Q}\left(\log\frac{Q}{\lambda}\right)\frac{d\lambda}{\lambda^2} &= \left(\frac{1}{\lambda}-\frac{1}{\lambda}\log\frac{Q}{\lambda}\right)\Big|_{Q_1}^{Q} = \frac{1}{Q_1}\log\frac{Q}{Q_1}+\frac{1}{Q}-\frac{1}{Q_1},\\
\int_{Q_1}^{Q}\left(\log\frac{Q}{\lambda}\right)\frac{d\lambda}{\lambda^{\frac12}} &= \left(2\lambda^{\frac12}\left(2+\log\frac{Q}{\lambda}\right)\right)\Big|_{Q_1}^{Q} = 4(Q^{\frac12}-Q_1^{\frac12}) - 2Q_1^{\frac12}\log\frac{Q}{Q_1},\\
\int_{Q_1}^{Q}\left(\log\frac{Q}{\lambda}\right)d\lambda &= \left(\lambda\log\frac{Q}{\lambda}+\lambda\right)\Big|_{Q_1}^{Q} = Q-Q_1-Q_1\log\frac{Q}{Q_1},\\
\int_{Q_1}^{Q}\left(\log\frac{Q}{\lambda}\right)\frac{d\lambda}{\lambda} &= \left(\frac12\left(\log\frac{Q}{\lambda}\right)^2\right)\Big|_{Q_1}^{Q} = -\frac12\left(\log\frac{Q}{Q_1}\right)^2
\end{aligned}$$

and eliminating the smaller terms we end up with

$$\sum_{\substack{q\le Q\\ l(q)>Q_1}} E^*(x,q) \ll Q(\log Qx)^2 + \left(\frac{x}{Q_1}\log x + Qx^{\frac12} + x^{\frac{13}{14}+\varepsilon}(\log x)^2\right)(\log x)^2.$$

5.4 Proof of Corollary 2

To prove Corollary 2 we use Proposition 2 and Siegel-Walfisz theorem. In fact we need to bound the sum over small q in (31). Siegel-Walfisz theorem gives

$$\max_{y\le x}|\psi(y,\chi)| \ll xe^{-c\sqrt{\log x}}$$

and summing it over all $q \leq (\log x)^A = Q_1$, Thus

$$\sum_{q \leq Q_1} \frac{1}{\varphi(q)} \sum_{\chi \bmod q}^{*} \max_{y \leq x} |\psi'(y, \chi)| \ll (\log x)^A x e^{-c\sqrt{\log x}} \ll \frac{x}{(\log x)^A} \ll x^{\frac{1}{2}} Q (\log x)^2.$$

5.5 Remarks

For an additional cancellation in type II sums one can also use the large sieve inequality with prime support.

Lemma (Large sieve inequality supported on primes)**.** *Let a_n be the sequence of complex numbers with the property that $a_n = 0$ if n has a prime divisor less than or equal to Q. Then*

$$\sum_{s\leq Q}\log\frac{Q}{s}\sideset{}{^*}\sum_{\chi(\mathrm{mod}\, s)}\left|\sum_{n=n_0}^{N}a_n\chi(n)\right|^2\leq (Q^2+(N-n_0+1))\sum_{n=n_0}^{N}|a_n|^2.$$

For the proof see [14, Theorem 9.11]. Using that one can derive the version needed for our purposes, analogous to Lemma 11 with

$$\sum_{s\leq Q}\log\frac{Q}{s}\sideset{}{^*}\sum_{\chi(\mathrm{mod}\, s)}\max_{y}\left|\sum_{m=m_0}^{M}\sum_{\substack{n=n_0\\ mn\leq y}}^{N}a_mb_n\chi(mn)\right|$$

instead of

$$\sum_{q\leq Q}\frac{q}{\varphi(q)}\sideset{}{^*}\sum_{\chi(\mathrm{mod}\, s)}\max_{y}\left|\sum_{m=m_0}^{M}\sum_{\substack{n=n_0\\ mn\leq y}}^{N}a_mb_n\chi(mn)\right|.$$

(One needs to soften the condition $Q\leq x^{\frac{1}{2}}$ to $Q\leq x^{\frac{1}{2}-\varepsilon}$ and allows the constant in $\ll$ to depend on ε.) Going further on applying this technique seems to be complicated, since one needs to get some cancellation in type I sums.

References

[1] A. Akbary and K. Hambrook. A variant of the Bombieri-Vinogradov theorem with explicit constants and applications. *Mathematics of Computation*, 84(294):1901–1932, 2014.

[2] M. Barban and P. Vekhov. An extremal problem. *Trudy Moscov. Mat. Obsc.*, 18:83–90, 1968.

[3] M. Bhargava, A. Shankar, T. Taniguchi, F. Thorne, J. Tsimerman, and Y. Zhao. Bounds on 2-torsion in class groups of number fields and integral points on elliptic curves. *preprint arXiv:1701.02458v1*, pages 1–12, 2017.

[4] E. Bombieri and J. Pila. The number of integral points on arcs and ovals. *Duke Mathematical Journal*, 3(59):337–357, 1926.

[5] A. Brumer. The average rank of elliptic curves I. *Inventiones Mathematicae*, 472(1):445–472, 1992.

[6] J. Cilleruelo, M.-C. Chang, M. Z. Garaev, J. Hernandez, I. E. Shparlinski, and A. Zumalacárregui. Points on Curves in Small Boxes and Applications. *Michigan Math. J.*, 63:503–534, 2014.

[7] J. Cilleruelo and I. E. Shparlinski. Concentration of points on curves in finite fields. *Monatshefte fur Mathematik*, 171(3-4):315–327, 2013.

[8] J. Cilleruelo, I. E. Shparlinski, and A. Zumalacárregui. Isomorphism classes of elliptic curves over a finite field in some thin families. *Math. Res. Lett.*, 19(02):1–9, 2012.

[9] A. C. Cojocaru and R. Murty. *An introduction to sieve methods and their applications*. London Mathematical Society Student Texts, 66. Cambridge University Press, Cambridge, Cambridge, 2006.

[10] H. Daboussi. Effective estimates of exponential sums over primes. *Analytic number theory, Vol. 1 (Allerton Park, IL, 1995)*, 138 of Pro(Birkhäuser Boston, Boston, MA):231–244, 1996.

[11] H. Daboussi and J. Rivat. Explicit upper bounds for exponential sums over primes. *Mathematics of Computation*, 70(233):431–447, 2000.

[12] H. Davenport. *Multiplicative Number Theory*. Springer-Verlag, 1980.

[13] F. Dress, H. Iwaniec, and G. Tenenbaum. Sur une somme liee a la fonction de Mobius. *J. Reine Angew. Math.*, 340:53–58, 1983.

[14] J. Friedlander and H. Iwaniec. *Opera de Cribro*. American Mathematical Society, 2010.

[15] D. Frolenkov. A numerically explicit version of the Pólya–Vinogradov inequality. *Moscow Journal of Combinatorics and Number Theory*, 1(3):234–250, 2011.

[16] P. X. Gallagher. A large sieve density estimate near sigma=1. *Inventiones Mathematicae*, 11(4):329–339, 1970.

[17] S. Graham. An asymptotic extimate related to Selberg's sieve. *Journal of Number Theory*, 10:83–94, 1978.

[18] R. Gross and J. H. Silverman. S-integer points on elliptic curves. *Pacific journal of mathematics*, 167(2):263–288, 1995.

[19] L. Hajdu and T. Herendi. Explicit Bounds for the Solutions of Elliptic Equations with Rational Coefficients. *Journal of Symbolic Computation*, 25:361–366, 1998.

[20] D. R. Heath-Brown. Prime numbers in short intervals and a generalized Vaughan identity. *Canad. J. Math.*, XXXIV(6):1365–1377, 1982.

[21] D. R. Heath-Brown. The Density of Rational Points on Curves and Surfaces. *Annals of Mathematics*, 155(2):553–595, 2002.

[22] H. Helfgott. *The ternary Goldbach problem*. Princeton University Press, submitted, preprint available arXiv:1501.05438.

[23] H. Helfgott. On the square-free sieve. *Acta Arithmetica*, 115(4):349–402, 2004.

[24] H. Helfgott and A. Venkatesh. Integral points on elliptic curves and 3-torsion in class groups. *J. Amer. Math. Soc.*, 19(3):527–550, 2006.

[25] H. Helfgott and A. Venkatesh. How small must ill-distributed sets be? *Analytic number theory*, 2:224–234, 2009.

[26] M. Hindry and J. H. Silverman. The canonical height and integral points on elliptic curves. *Inventiones Mathematicae*, 93(2):419–450, 1988.

[27] G. Kabatiansky and V. Levenstein. On bounds for packings on a sphere and in soace. *Problemy Peredachi Informacii*, 14(1):3–25, 1978.

[28] E. Landau. Über Ideale und Primideale in Idealklassen. *Mathematische Zeitschrift*, 2(1-2):52–154, 1918.

[29] S. Lang. *Elliptic curves: Diophantine analysis*, volume 231. Springer-Verlag, 1978.

[30] H. Lenstra and C. Pomerance. Primality testing with Gaussian periods. *preprint*, pages 1–47, 2016.

[31] J. Milne. On a conjecture of Artin and Tate. *Annals of Mathematics*, 102:517–533, 1975.

[32] J. Milne. *Elliptic Curves*. BookSurge Publishers, 2006.

[33] H. L. Montgomery. The analytic principle of the large Sieve. *Bulletin of the American Mathematical Society*, 84(4):547–567, 1978.

[34] A. Pacheco. Integral points on elliptic curves ove function fields of positive characteristic. *Bull. Austral. Math. Soc.*, 58:353–357, 1998.

[35] A. Page. On the number of primes in an arithmetic progression. *Proc. London Math. Soc. S2-39*, 1:116, 1933.

[36] P. S. Park. The Bombieri–Vinogradov theorem. *Expository, preprint available http://web.math.princeton.edu/˜pspark/papers/bv.pdf*, pages 1–33, 2016.

[37] C. Pomerance. Remarks on the Polya-Vinogradov inequality. *Integers*, 4:531–542, 2010.

[38] B. Poonen. Gonality of modular curves un characteristics p. *Math. Res. Lett.*, 14(4):691–701, 2007.

[39] J. H. Silverman. A lower bound for the canonical height on elliptic curves. *Duke Mathematical Journal*, 48(3):633–648, 1981.

[40] J. H. Silverman. Lower bounds for height functions. *Duke Mathematical Journal*, 51(2):395–403, 1984.

[41] J. H. Silverman. *Advanced topics in the arithmetic of elliptic curves*. Graduate Texts in Mathematics, 151. Springer-Verlag, New York, 1994.

[42] H. Swinnerton-Dyer. The Number of Lattice Points on a Convex Curve. *Journal of Number Theory*, 135(January 1971):128–135, 1974.

[43] J. Tate. On the conjectures of Birch and Swinnerton- Dyer and a geometric analog. *Séminaire N. Bourbaki*, 306:415–440.

[44] G. Tenenbaum. *Introduction to analytic and probabilistic number theory*. Graduate Studies in Mathematics, 163. American Mathematical Society, Providence, RI, 2015., third edit edition, 1995.

[45] N. Timofeev. The Vinogradov-Bombieri theorem. *Mat. Zametki*, 38(6):801–809, 1985.

[46] D. Ulmer. *Elliptic curves over function fields*. Arithmetic of L-functions; 211-280, IAS/Park City Math. Ser., 18, Amer. Math. Soc., Providence, RI, 2011.

[47] R. C. Vaughan. Mean value theorems in prime number theory. *J. London Math. Soc.*, 10:153–162, 1975.

[48] R. C. Vaughan. Sommes trigonométriques sur les nombres premiers. *C.R. Acad. Sci. Paris*, 285(16):A981–A983, 1977.

[49] I. M. Vinogradov. *The method of triginometrical sums in the theory of numbers*. Interscience, New York, 1954.

[50] F. Voloch. Explicit p-descent for elliptic curves in characteristic p. *Compositio Math*, 74(3):247–258, 1990.