

M. KLAZAR

On the recent result of Green and Tao about
arbitrarily long arithmetic progressions
~~of~~ primes

B. Green (1977) & T. Tao (1975), Apr 8, 2004,
Artiv: $\forall \exists$ primes $p_1 < p_2 < \dots < p_k$ s.t.
 $p_2 - p_1 = p_3 - p_2 = \dots = p_k - p_{k-1} (= \Delta)$

Related results

Dirichlet (1837): $\forall a, m \in \mathbb{N}, (a, m) = 1 \exists p \equiv a \pmod{m}$
Mordell (1930): $\exists k$ s.t. $n = p_1 + p_2 + \dots + p_k$,
 $k \leq 2$, has solution $\forall n \geq 2$.

Budorragob (1937): $2n+1 = p_1 + p_2 + p_3$ has so-
lution $\forall n > n_0$.

Van der Corput (1939): $\exists \infty$ many triples

$$\begin{array}{ccc} \leftrightarrow & \leftrightarrow & \leftrightarrow \\ p_1 & p_2 & p_3 \end{array}$$

Heath-Brown (1981): $\exists \infty$ many quadruples

$$\begin{array}{cccc} \leftrightarrow & \leftrightarrow & \leftrightarrow & \leftrightarrow \\ n_1 & p_2 & p_3 & n_4 \end{array} \text{ where } \Omega(n_1 n_4) \leq 3.$$

Szemerédi's theorem (1975)

$\forall \delta > 0, k \geq 3 \exists N_0(\delta, k) : N \geq N_0(\delta, k) \wedge x \in \mathbb{Z}_N$
with $|x| \geq \delta N$, then $x \supset \text{A.P. of length } k$.
[\mathbb{Z}_N is $\mathbb{Z}/N\mathbb{Z}$, N big prime]

The proof of Green and Tao consists of 2 steps

① Szemerédi's thm. \Rightarrow Relative Sz. thm.

Ergodic-combinatorial part :

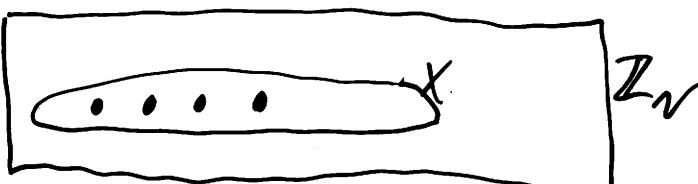
RST: $\delta > 0, k \geq 3, \tilde{x} \in \mathbb{Z}_N$ pseudorandom, $x \subset \tilde{x}, \frac{|x|}{|\tilde{x}|} \geq \delta$.
Then $x \supset \text{A.P. of length } k$, for N big enough.

② Number-theoretic part

$\mathbb{Z}_N \supset \mathcal{P}$ -primes ($< N$). \exists pseudor. $\tilde{\mathcal{P}} \supset \mathcal{P}$ s.t. $\frac{|\mathcal{P}|}{|\tilde{\mathcal{P}}|} \geq \delta$.

① & ② $\Rightarrow \mathcal{P} \supset \text{A.P. of length } k$ for $k \geq 3$.

Sz. Thm.:



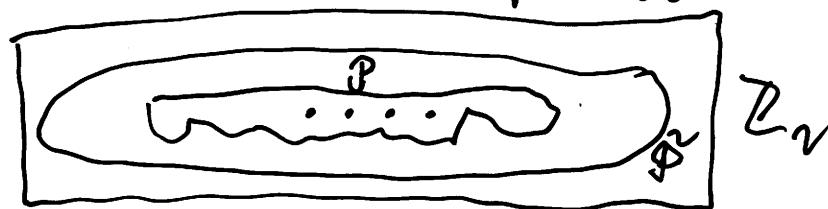
$$|x|/N \geq \delta$$

R. Sz. Thm.



$$|x|/|\tilde{x}| \geq \delta$$

Sz. Thm. for primes



$$|\mathcal{P}|/|\tilde{\mathcal{P}}| \geq \delta$$

In more details ... N -large prime, $\mathcal{O}(1)$, $\mathcal{O}(1)$ -
w.r.t. $N \rightarrow \infty$. 3

① Relative Szemerédi's theorem

A-finite set, $f: A \rightarrow \mathbb{R}$, $\mathbb{E}(f) = \mathbb{E}(f(x) | x \in A) :=$
 $= \frac{1}{|A|} \sum_{x \in A} f(x)$. $\nu: \mathbb{Z}_N \rightarrow \mathbb{R}^+$ measure: $\mathbb{E}(\nu) = 1 + \mathcal{O}(\epsilon)$

(Thm. 2.3)

Szemerédi's Thm.: $0 < \delta \leq 1$, $q \geq 3$, $f: \mathbb{Z}_N \rightarrow \mathbb{R}^+$ s.t.
 $0 \leq f(x) \leq 1 \quad \forall x \in \mathbb{Z}_N \quad \& \quad \mathbb{E}(f) \geq \delta$.

Then $\mathbb{E}\left(\prod_{i=0}^{q-1} f(x+ir) \mid x, r \in \mathbb{Z}_N\right) \geq c(q, \delta) - \mathcal{O}(1)$.

> 0

Relative Szemerédi's thm. (Thm. 3.5): $0 < \delta \leq 1$, $q \geq 3$,
 $\nu: \mathbb{Z}_N \rightarrow \mathbb{R}^+$ q -pseudor. measure, $f: \mathbb{Z}_N \rightarrow \mathbb{R}^+$ s. f.
 $0 \leq f(x) \leq \nu(x) \quad \forall x \in \mathbb{Z}_N \quad \& \quad \mathbb{E}(f) \geq \delta$.

Then $\mathbb{E}\left(\prod_{i=0}^{q-1} f(x+ir) \mid x, r \in \mathbb{Z}_N\right) \geq c(q, \delta) - \mathcal{O}(1)$.

[$c(q, \delta) > 0$ the same in both thm's. $\mathcal{O}(1) = \sigma_{q, \delta}(1)$.]

So constant 1 was "just" replaced by any
 q -pseudorandom measure $\nu(x)$.

What is q -pseudorandomness?

(4)

A measure $V: \mathbb{Z}_N \rightarrow \mathbb{R}^+$ is \mathbb{Q} -pseudorandom if it satisfies 2 conditions.

a) (lin. forms condition)

$$L = \boxed{\in \mathbb{Q}^{m \times t}} \quad \begin{matrix} m \leq 2^{k-1}, L_{ij} = \frac{a}{b} \in \mathbb{Q}, |a|, |b| \leq \mathbb{Q} \\ t \leq 3k-4 \end{matrix} \quad b_1, b_2, \dots, b_m \in \mathbb{Z}_N$$

$$\psi_i(\bar{x}) = \sum_{j=1}^t L_{ij} x_j + b_i, \quad i=1, 2, \dots, m.$$

L nondegenerate (no 0-row, no row is a multiple of another row). Then

$$\mathbb{E}\left(\prod_{i=1}^m V(\psi_i(\bar{x})) \mid \bar{x} \in \mathbb{Z}_N^t\right) = 1 + o_{\mathbb{Q}}(1).$$

b) (correlation condition)

If $m, m \leq 2^{k-1}$, \exists weight function $\tilde{T}_m: \mathbb{Z}_N \rightarrow \mathbb{R}^+$ s.t. $\mathbb{E}(\tilde{T}_m^q) = O_{\mathbb{Q}, q}(1)$ $\forall q \in \mathbb{N}$ and

$$\mathbb{E}\left(\prod_{i=1}^m V(x+h_i) \mid x \in \mathbb{Z}_N\right) \leq \sum_{1 \leq i < j \leq m} \tilde{T}_m(h_i - h_j)$$

\exists m -tuple $h_1, h_2, \dots, h_m \in \mathbb{Z}_N$ (not necessarily distinct).

[In b) $L = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}$ is degenerate, $b_i = h_i$.]

(2) Enveloping Primes by a pseudor. seq; with
 $\gg 0$ relative density of primes.

$\Lambda : N, \mathbb{Z}_N \rightarrow \mathbb{R}^+$, $\Lambda(n) = \sum \log p \dots n = p^r$

von Mangoldt function $\begin{cases} 0 & \dots n \neq p^r \end{cases}$

[Case $r \geq 2$ will be neglected]

Recall that $\Lambda(n) = \sum_{d|n} \mu(d) \log \frac{n}{d}$

[Möbius inverse of $\log n = \sum_{d|n} \Lambda(d)$, μ is M. function
 $\mu(n) = \begin{cases} (-1)^r & \dots n = p_1 p_2 \dots p_r \\ 0 & \dots n \text{ is not } \square\text{-free} \end{cases}$

PNT $\rightarrow E(\Lambda) = 1 + o(1)$ and Λ is supported by
 primes (we can neglect $p^r, r \geq 2$).

Suppose we have a \mathbb{Q} -pseudor. measure ν s.t.
 $\nu(n) \geq c\epsilon q$. $\Lambda(n) \neq 0 \in \mathbb{Z}_N$. Then we have the
 main result. Unfortunately, no such ν exists.

[$q \in \mathbb{N}$, Λ concentrated only on those \pmod{q}
 classes, for which $(q, q) = 1$ but ν must be evenly
 distributed among all $q \pmod{q}$ classes. ~~and~~ And
 $\liminf_{q \rightarrow \infty} \frac{\nu(q)}{q} = 0$.] G.H.T. get around this by the U-trick

$$w = w(N) \rightarrow \infty \text{ slowly with } N \rightarrow \infty. \quad (6)$$

$$W = W(N) = \prod_{p \leq w(N)} P \quad \left| \begin{array}{l} \tilde{\chi}(n) = \begin{cases} \frac{\varphi(n)}{n} \log(kn+1) \\ \dots \text{ if } kn+1 \text{ is prime} \\ 0 \dots \text{ otherwise} \end{cases} \end{array} \right.$$

A.P. in n 's \rightarrow A.P. in the values $kn+1$.

Enveloping primes (Prop. 8.1.) $\varepsilon_2 = \frac{1}{2^2(2+4)!}$

N prime and big - then \exists 2-pseudor. measure V :
 $\mathbb{Z}_N \rightarrow \mathbb{R}^+$ s.t.

$$V(n) \geq \frac{1}{2^2 2^{2+5}} \tilde{\chi}(n) \text{ holds for}$$

every $n \in [\varepsilon_2 N, 2\varepsilon_2 N]$.

Proof of the main result.

Set $f(n) = \begin{cases} 2^{-1} 2^{-2-5} \tilde{\chi}(n) & n \in [\dots] \\ 0 & \text{else.} \end{cases}$

$$\mathbb{E}(f) = N^{-1} 2^{-1} 2^{-2-5} \sum_{n \in [\dots]} \tilde{\chi}(n) = \dots = \frac{\varepsilon_2}{2^2 2^{2+5}} (1 + o(1)) = o(1)$$

[by PNT for ar. progressions]. By Prop. 8.1 and Thm. 3.5, $\mathbb{E}\left(\prod_{i=0}^{2-1} f(x+i\tau) \mid x, \tau \in \mathbb{Z}_N\right) \geq c(2, \delta) - o(1)$

> 0 for big N - we get a genuine A.P. of length 2 of primes in \mathbb{N} .

$[\varepsilon_2 < \frac{1}{2} \cdot \text{Also, case } \tau = 0 \text{ contributes only } O\left(\frac{\log^2 N}{N}\right) = o(1)]$

How is the λ -pseudo random measure $V(n)$
that majorizes $C(n)$. $\tilde{A}(n)$ defined?

$$\Lambda(n) = \sum_{d|n} \mu(d) \log \frac{n}{d} = \sum_{d|n} \mu(d) \log_+ \frac{n}{d} \text{ where } \log_+ x = \max(\log x, 0).$$

$R > 0$ parameter [$R = N^{\text{smaller}}$]

$$\Lambda_R(n) := \sum_{\substack{d|n \\ d \leq R}} \mu(d) \log \frac{R}{d} = \sum_{d|n} \mu(d) \log_+ \frac{R}{d}.$$

(Goldston - Yildirim)

Definition of V (Def. 8.3) $R := N^{2^{-1}2^{-8-4}}$ and ε_R as

before ($\varepsilon_R = \frac{1}{2^8(8+4)!}$). We define $V: \mathbb{Z}_N \rightarrow \mathbb{Q}^+$ by

$$V(n) = \prod_{k=1}^r \frac{\varphi(u_k)}{u_k} \cdot \frac{1}{\log R} \cdot \Lambda_R(u_{n+1})^2 \dots u \in [\varepsilon_R N, 2\varepsilon_R N]$$

1 ... otherwise

It is shown that $V(n) \geq \frac{1}{2^{2^{8+5}}} \tilde{A}(n)$ (Lemma 8.4)

$E(V) = 1 + o(1)$ (Lemma 8.7) and that V is λ -pseudo-random (Prop. 8.8 and Prop. 8.10). The last two results are based on asymptotics for expressions involving $\Lambda_R(n)$, derived by method due to Goldston and Yildirim.

Further details

18

① Relative Szemerédi's theorem (Thm. 3.5)

Its proof uses 3 results : Szemerédi's thm.
 Generalized von Neumann thm. (Thm. 2.3)
 Generalized Koopman-von N. thm. (Prop. 7.1).

$f: \mathbb{Z}_N \rightarrow \mathbb{R}$, $\|f\|_{U^d}$ - Gowers' uniformity norm
Generalized von N. thm. (Prop. 5.3)

$V: \mathbb{Z}_N \rightarrow \mathbb{R}^+$ q -pseudor. measure, $f_0, \dots, f_{q-1}: \mathbb{Z}_N \rightarrow \mathbb{R}$
 are s.t. $|f_i(x)| \leq V(x) + 1 \quad \forall x \in \mathbb{Z}_N, i = 0 \dots q-1$.

$$[\Rightarrow \leq V(x)]$$

Then $\mathbb{E} \left(\prod_{i=0}^{q-1} f_i(x+ir) \mid x, r \in \mathbb{Z}_N \right) = O(\min_{i=0 \dots q-1} \|f_i\|_{U^{q+1}}) + o(1)$.

Definition of $\|\cdot\|_{U^d}$.

$f: \mathbb{Z}_N \rightarrow \mathbb{R}$, $w \in Q_d = \{0, 1\}^d$, $|w| = w_1 + \dots + w_d$. For
 $h \in \mathbb{Z}_N^d$ we set $w \cdot h = w_1 h_1 + \dots + w_d h_d$.

$\{x + w \cdot h \mid w \in Q_d\}$ is d -dim cube in \mathbb{Z}_N .

$$\|f\|_{U^d} := \left(\mathbb{E}_{w \in Q_d} \left(\prod_{h \in \mathbb{Z}_N^d} f(x + w \cdot h) \mid x \in \mathbb{Z}_N \right) \right)^{1/2^d}$$

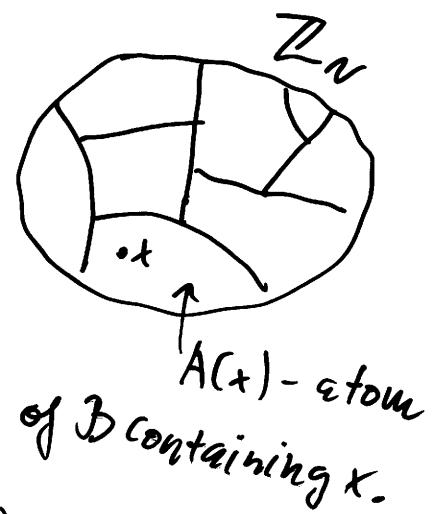
Note that $\|f\|_{U^1} = (\mathbb{E}(f)^2)^{1/2} = |\mathbb{E}(f)|$ is just semi-norm. For $d \geq 2$ $\|\cdot\|_{U^d}$ is a norm.

σ -algebra \mathcal{B} in \mathbb{Z}_N : $\mathcal{B} \subset \wp(\mathbb{Z}_N)$ closed to \cup, \cap, \setminus ; $\emptyset, \mathbb{Z}_N \in \mathcal{B}$. Atoms of \mathcal{B} : minimal $\neq \emptyset$ elements; form a partition of \mathbb{Z}_N :

$$f: \mathbb{Z}_N \rightarrow \mathbb{R}, E(f|\mathcal{B}): \mathbb{Z}_N \rightarrow \mathbb{R}$$

$$\text{given by } E(f|\mathcal{B})(x) = \frac{1}{|A(x)|} \sum_{y \in A(x)} f(y)$$

- Conditional expectation of f w.r.t. \mathcal{B} .



Generalized Koopman-von N. thm. (Prop. 7.1)

$V: \mathbb{Z}_N \rightarrow \mathbb{R}^+$ 2-pseudor. measure, $f: \mathbb{Z}_N \rightarrow \mathbb{R}$ s.t.

$0 \leq f(x) \leq V(x) \quad \forall x \in \mathbb{Z}_N, 0 < \varepsilon < 1, N$ big (dep. on ε). Then $\exists \sigma$ -alg. $\mathcal{B}, \Omega \in \mathcal{B}$ s.t.

$E(V \cdot \chi_\Omega) = \varepsilon(1) \dots$ i.e., Ω is small (exceptional) set
 $\|(1 - \chi_\Omega)E(V - 1|\mathcal{B})\|_{L^\infty} = \varepsilon(1) \dots$ i.e., \sim .

$\|(1 - \chi_\Omega)(f - E(f|\mathcal{B}))\|_{L^{2/1}} \leq \varepsilon^{1/2} \dots$ i.e., \sim .

Proof of Prop. 7.1 takes cca 12 pages and uses, e.g., Weierstrass approximation thm. Proof of Prop. 5.3 takes cca 4 pages and uses Cauchy-Schwarz ineq. and 2-pseudorandomness of V .

Thm. 2.3 is a black box (proof takes 50, resp. 20, pages elsewhere).

To prove Relative Sz. thm. (Thm. 3.5) is not easy (1/2 page). We have f , σ and V as in Thm. 3.5. Take B and ω from Prop. 7.1 and write

$$f = \underbrace{(f - E(f|B))}_{g} + \underbrace{E(f|B)}_{h}. \text{ Express}$$

$$E\left(\prod_{i=0}^{q-1} f(x+ir) \mid x, r \in \mathbb{Z}_N\right) \text{ as } \begin{matrix} \text{the} \\ \text{main term + error:} \end{matrix}$$

$$= g(\omega) + h(\omega)$$

$$= E\left(\prod_{i=0}^{q-1} h(x+ir) \mid \omega\right) + \sum_{\substack{\text{2}^{q-1} \text{ terms} \\ i=0}} E\left(\prod_{i=0}^{q-1} g(x+ir) \mid \omega\right)$$

M E

at least once g

We work outside the negligible Ω . $M \geq c(q, \delta) > 0$ by Sz. thm. (Thm. 2.3). ($E(V|B)$ behaves like $\varepsilon^{-\delta_\varepsilon(1)}$ atoms outside Ω)

$E = O(\varepsilon^{q/2})$ by Generalized von N. thm. (Prop. 5.3).

So $E(\omega) \geq c(q, \delta) - O(\varepsilon^{q/2}) - \delta_\varepsilon(1)$.

② Enveloping Primes

71

We need to prove that $\nu(n) = \begin{cases} \frac{\varphi(\kappa)}{\kappa} \cdot \frac{\Lambda_R(Wn+1)^2}{\log R} & \dots \text{for } n \in [\varepsilon_\alpha N, 2\varepsilon_\alpha N] \\ 1 \dots \text{otherwise} & \end{cases}$

is α -pseudorandom ($R = N^{2^{-1}2^{-\alpha-4}}$, $\Lambda_R(n) = \sum_{d|n} \mu(d) \log \frac{R}{d}$)
 $\varepsilon_\alpha = \frac{1}{2^\alpha (\alpha+4)!}$).

The lin. forms condition on V is proved by the following asymptotics due to Goldston and Yildirim.

Prop. 8.5

$$L = \boxed{ \in \mathbb{Z}^{m \times t} } \quad \begin{matrix} m \\ \uparrow \\ \leftarrow \end{matrix} \quad \begin{matrix} \text{- nondegenerate matrix (cf. 1.8-condition)} \\ |L_{ij}| \leq \frac{1}{2} \sqrt{w(N)}, b_1, \dots, b_m \in \mathbb{Z} \end{matrix}$$

$$\psi_i(\bar{x}) = \sum_{j=1}^t L_{ij} x_j + b_i, i = 1 \dots m \quad \begin{matrix} \text{intervals} \\ \swarrow \end{matrix}$$

$B = I_1 \times I_2 \times \dots \times I_t \subset \mathbb{R}^t$, $|I_j| \geq R^{10m}$ $\forall j$. Then (if $w(N) \rightarrow \infty$ as $N \rightarrow \infty$ slowly enough)

$$\mathbb{E} \left(\prod_{i=1}^m \Lambda_R(W\psi_i(\bar{x}) + 1)^2 \mid \bar{x} \in B \right) \approx (1 + o_m(1)) \left(\frac{W \log R}{\varphi(W)} \right)^m$$

Similar Prop. 8.6 takes care of the correlation condition; it is proved by methods due to Goldston and Yildirim as well.

Let us review the proof of Prop. 8.5.

$B = \prod_{j=1}^t I_j \subset \mathbb{R}^t$, $|I_j| \geq R^{2m}$. We seek the asymptotics of $\mathbb{E} \left(\prod_{i=1}^m \chi_R(W\psi_i(\bar{x})+1)^2 \mid \bar{x} \in B \right)$. (*)

(*) is transformed (upto a small error) in the integral form

$$\frac{1}{(2\pi i)^m} \int_{\Gamma_1} \dots \int_{\Gamma_m} F(z, u) \prod_{j=1}^m \frac{R^{z_j + u_j}}{z_j^2 u_j^2} dz_j du_j \quad (**)$$

where

$$z \in \mathbb{C}^m \\ z = z_1, \dots, z_m ; u \in \mathbb{C}^m \\ u = u_1, \dots, u_m ; d \in \mathbb{N}^m \\ d = d_1, \dots, d_m ; e \in \mathbb{N}^m$$

$$F(z, u) = \sum_{\substack{d_1, \dots, d_m \\ e_1, \dots, e_m \in \mathbb{N}}} \left(\prod_{j=1}^m \frac{\mu(d_j) \mu(e_j)}{d_j^{z_j} \cancel{u_j^{e_j}}} \right) \prod_p H(p, d, e)$$

$$H(p, d, e) = \mathbb{E} \left(\prod_{j=1}^m \chi_{W\psi_j(\bar{x})+1 \equiv 0 \pmod p} \mid \bar{x} \in \mathbb{Z}_p^t \right) =$$

$$= \frac{1}{|\mathbb{Z}_p^t|} \#\{ \bar{x} \in \mathbb{Z}_p^t : p \nmid d_j e_j \Rightarrow W\psi_j(\bar{x})+1 \equiv 0 \pmod p \}_{j=1 \dots m},$$

and $\Gamma_1 = -\frac{i}{2} \log R + \frac{1}{2} \log R$ This is derived by using definition of χ_R , changing summation order and using the formula $\frac{1}{2\pi i} \int_{\Gamma_1} \frac{z^x}{z^2} dz = \log$ (known from Perron's Σ or theory of $\zeta(s)$).

$D_\sigma^m = \{z_j, u_j : -\sigma < \operatorname{Re}(z_j), \operatorname{Re}(u_j) < 100, j=1 \dots m\}$, for $\sigma > 0$
 (13)

$G = G(z, u)$ holom. in D_σ^m we set $V(G, \delta, D)$ to be
 the largest $\sup_{z, u \in D} |\tilde{G}(z, u)|$ where \tilde{G} = some ∂G of
 order $\leq \frac{m}{2}$.

$F = F(z, u)$, appearing in (**), factorizes as $F =$
 $= G_1 G_2 G_3$ where $G_3(z, u) = \prod_{j=1}^m \frac{\zeta(1+z_j+u_j)}{\zeta(1+z_j)\zeta(1+u_j)}$ ($\sum_{n=1}^m n^{-s}$)

G_1, G_2 are holom. in $\underbrace{D_{1/6m}}_D$ and satisfy

$V(G_1, m, D) \leq \Omega_m(1)$, $V(G_2, m, D) \leq w(n)^{\Omega_m(w(n))}$
 $G_1(\bar{z}, \bar{u}) = 1 + o_m(1)$, $G_2(\bar{z}, \bar{u}) = (w/\varphi(w))^u$. (Lemma 9.3)
 Thus (**) can be evaluated by

Lemma 9.4 (6.84.) $R > 0$, $G = G_R(z, u)$ holom. in D_R^m

for some $\sigma > 0$, $V(G, m, D) = \exp(\Omega_{m, \sigma}(\log^{2/3} R))$.

Then $\frac{1}{(2\pi i)^m} \int_{\Gamma_1} \dots \int_{\Gamma_m} G(z, u) \prod_{j=1}^m \frac{\zeta(1+z_j+u_j)}{\zeta(1+z_j)\zeta(1+u_j)} \frac{dz_j+u_j}{R} dz_j du_j$

$= G(\bar{z}, \bar{u}) \log R + \sum_{j=1}^m \Omega_{m, \sigma} (V(G, j, D) \cdot \log^{m-j} R) + \Omega_{m, \sigma} (e^{-\delta \sqrt{\log R}})$

for some $\delta = \delta(m) > 0$. Asymptotics of (*) = (**) follows by
 setting $G = G_1 G_2$ and $D = 1/6m$.

Lemma 9.4 ~~takes care also~~ is also used in the proof of Prop. 8.6. The proof of Lemma 9.4 takes 4 pages in the Appendix. It uses contour deformation, residues and 0-free region for $\xi(s)$.

Want even more details? Read and ponder the article!

On May 26, 2004, R.F. Arenstorf posted in Artiv the preprint "There are infinitely many prime twins".... (i.e., $\exists \infty$ many $p_1 < p_2, p_2 - p_1 = 2$)

Last news: the first 8 primes in an A.P. appear before the number

$$2^{2^{2^{2^{2^{2^{2^{2^{2^{100}}}}}}}}$$

(June 2, home page of T.Tao).