A counterexample relating exponential sums and discrepancy

Daniel Miller

December 22, 2016

For a prime p, let

$$T_p = \left\{ \frac{a}{2\sqrt{p}} : a \in \mathbf{Z}, |a| \leqslant 2\sqrt{p} \right\}$$
$$\Theta_p = \cos^{-1}\left(T_p\right).$$

Since applying continuous increasing functions preserves discrepancy, we have:

$$\operatorname{disc}(T_p, \mathrm{Leb}) \ll p^{-1/2}$$
$$\operatorname{disc}\left(\Theta_p, \frac{1}{2}\sin(t)\,\mathrm{d}t\right) \ll p^{-1/2}.$$

We claim that starting with $\theta_2 \in \Theta_2$, we can choose θ_p such that we preserve the inequalities:

$$\frac{1}{4\log x} \leqslant \operatorname{disc}(\{\theta_p\}_{p\leqslant x}) \leqslant \frac{4}{\log x}$$
$$\left| \sum_{p\leqslant x} U_1(\theta_p) \right| \leqslant 2\sqrt{x}$$

Recall that

$$U_1(\theta) = \frac{\sin(2\theta)}{\sin\theta}.$$

We can run this for all $p \leqslant 10^5$. Recall that $\pi(10^5) \approx 10000$.

Here is what we get:

Conjecture 1. *There exists a sequence of $\theta_p \in \Theta_p$ such that the following identities always hold:*

$$\frac{1}{4\log x} \leqslant \operatorname{disc}(\{\theta_p\}_{p\leqslant x}) \leqslant \frac{4}{\log x}$$
$$\left| \sum_{p\leqslant x} U_1(\theta_p) \right| \leqslant 2\sqrt{x}.$$

Figure 1: Plot of $\sum_{p \leqslant x} U_1(\theta_p)$

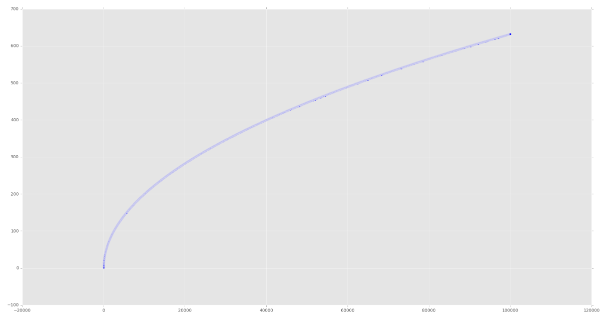
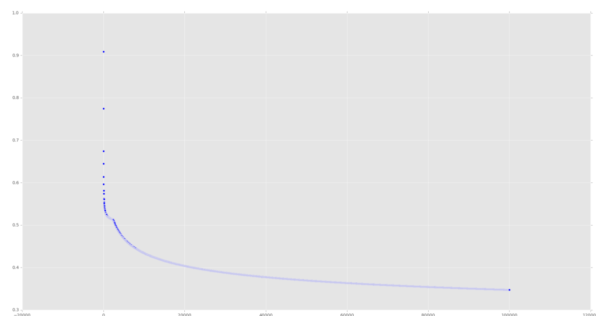


Figure 2: Plot of $\mathrm{disc}(\{\theta_p\}_{p \leqslant x})$



Next, choose $\bar{\rho}_l \colon G_{\mathbf{Q}} \twoheadrightarrow \mathrm{GL}_2(\mathbf{F}_l)$ to which we can apply Ramakrishna et. al.'s machinery. Define

$$\Theta_p(\bar{\rho}_l) = \left\{ \cos^{-1}\left(\frac{a}{2\sqrt{p}}\right) : a \in \mathbf{Z}, |a| \leqslant 2\sqrt{p}, a \equiv \mathrm{tr}\, \bar{\rho}_l(\mathrm{fr}_p) \pmod{l} \right\}.$$

Conjecture 2. *There exists a sequence of $\theta_p \in \Theta_p(\bar{\rho}_l)$ such that*

$$\mathrm{disc}(\{\theta_p\}_{p \leqslant x}) = \Omega\left(\frac{1}{\log x}\right)$$

$$\left| \sum_{p \leqslant x} U_1(\theta_p) \right| \ll \sqrt{x}.$$

Corollary 1. *There exists an (infinitely ramified) Galois representation $\rho_l \colon G_{\mathbf{Q}} \to \mathrm{GL}_2(\mathbf{Z}_l)$ such that if we set $a_p = \mathrm{tr}\, \rho_l(\mathrm{fr}_p)$, then*

1. $a_p \in \mathbf{Z}$

2. $|a_p| \leqslant 2\sqrt{p}$ *.*

3. The $\theta_p = \cos^{-1}\left(\frac{a_p}{2\sqrt{p}}\right)$ *satisfy*

$$\operatorname{disc}(\{\theta_p\}_{p\leqslant x}) = \Omega\left(\frac{1}{\log x}\right)$$

$$\left|\sum_{p\leqslant x} U_1(\theta_p)\right| \ll \sqrt{x}.$$

and hence $L(\rho_l, s)$ *satisfies the Riemann Hypothesis.*

1 Towards a proof

Let $\bar{\rho}_l \colon G_{\mathbf{Q}} \to \mathrm{GL}_2(\mathbf{F}_l)$ be a Galois representation. For each prime p, define

$$\Theta_p(l) = \left\{\cos^{-1}\left(\frac{a}{2\sqrt{p}}\right) : a \in \mathbf{Z}, |a| \leqslant 2\sqrt{p}, a \equiv \operatorname{tr}\bar{\rho}_l(\mathrm{fr}_p) \pmod{l}\right\}.$$

It is easy to check that

$$\operatorname{disc}\left(\Theta_p(l), \frac{1}{2}\sin(t)\,\mathrm{d}t\right) \ll lp^{-1/2}.$$

We are looking for a way to choose $\theta_p \in \Theta_p(l)$ such that

1. $\operatorname{disc}(\{\theta_p\}_{p\leqslant x})$ decays like $1/\log x$
2. $\left|\sum_{p\leqslant x} U_1(\theta_p)\right|$ grows like $\sqrt{x}$.

To do this, suppose we have chosen $\{\theta_q\}_{q<p}$. In choosing θ_p, we want to simultaneously move the discrepancy towards $1/\log p$, while making sure that the U_1-sum doesn't get too big.

(Fact: if $\{x_1, \ldots, x_N\}$ and $\{y_1, \ldots, y_N\}$ are two sequences, then

$$|\operatorname{disc}(\{x_1, \ldots, x_N\}) - \operatorname{disc}(\{y_1, \ldots, y_N\})| \leqslant 2\|x - y\|_\infty.$$

)

It's actually quite simple. Note that:

$$U_1(\theta) = \frac{\sin(2\theta)}{\sin\theta} = -U_1(\pi - \theta).$$

The basic idea is: set $\theta_3 \approx \pi - \theta_2$, $\theta_7 \approx \pi - \theta_5$, etc. and we can choose θ_2, θ_5 etc. arbitrarily, meaning good discrepancy, while the sum should approximately cancel out. First, since U_1 has bounded derivative, we know that

$$|U_1(\theta) - U_1(\varphi)| \ll |\theta - \varphi|$$

So, if $p_1 < p_2$ are sequential primes, we have

$$|\theta_{p_2} - (\pi - \theta_{p_1})| \ll p_1^{-1/2},$$

so

$$\begin{aligned}|U_1(\theta_{p_1}) + U_1(\theta_{p_2})| &\leqslant |U_1(\theta_{p_1}) - U_1(\pi - \theta_{p_1})| + |U_1(\pi - \theta_{p_1}) - U_1(\theta_{p_2})| \\ &\ll |\theta_{p_2} - (\pi - \theta_{p_1})| \\ &\ll p_1^{-1/2}.\end{aligned}$$

So,

$$\left|\sum_{p \leqslant x} U_1(\theta_p)\right| \ll \sum_{p \leqslant x} p^{-1/2} \ll \int_1^x t^{-1/2}\,\mathrm{d}t \ll \sqrt{x}.$$

(Same argument works for all U_{odd} because they all satisfy $U_{\mathrm{odd}}(\pi - \theta) = -U_{\mathrm{odd}}(\theta)$. In contrast, $U_{\mathrm{even}}(\pi - \theta) = U_{\mathrm{even}}(\theta)$.)

2 A legit proof!

Theorem 1. *Fix a prime l. Suppose we have chosen, for all primes p, some arbitrary residue class $\bar{a}_p \in \mathbf{F}_l$, and set*

$$\Theta_p(l) = \left\{\cos^{-1}\left(\frac{a}{2\sqrt{p}}\right) : a \in \mathbf{Z}, |a| \leqslant 2\sqrt{p}, a \equiv \bar{a}_p \pmod{l}\right\}.$$

Then there exists a choice of $\theta_p \in \Theta_p(l)$ such that

1. *The sequence $\{\theta_p\}$ is equidistributed with respect to the Sato–Tate measure $\frac{2}{\pi}\sin^2\theta\,\mathrm{d}\theta$.*
2. *The discrepancy* $\mathrm{disc}(\{\theta_p\}_{p \leqslant x}, \mathrm{ST}) \gg \frac{1}{\log x}$.
3. $\left|\sum_{p \leqslant x} U_{\mathrm{odd}}(\theta_p)\right| \ll \sqrt{x}$.

Proof. Enumerate the primes $p_1 < p_2 < \cdots$. We will choose $\theta_{p_{\mathrm{odd}}} \in [0, \pi/2)$ so that the discrepancy of the sequence $\{\theta_{p_{\mathrm{odd}}}\}$ behaves as required in that interval. We'll then set $\theta_{p_{2i}} \approx \pi - \theta_{p_{2i-1}}$.

Everything comes down to: if $p < q$ are sequential primes and we have already chosen θ_p, we need to be able to choose θ_q so that $|U_1(\theta_p) + U_1(\theta_q)| \ll p^{-1/2}$. Since $\frac{\mathrm{d}U_1}{\mathrm{d}\theta} = -2\sin(\theta)$, we have (roughly)

$$|U_1(\theta) - U_1(\varphi)| \ll \max(\theta, \varphi) \cdot |\theta - \varphi|$$

for $\theta, \varphi \in [0, \pi/2)$.

Start with $t_p = \frac{a_p}{2\sqrt{p}}$ and $t_q = \frac{a_q}{2\sqrt{q}}$. We can guarantee that $|t_p - (\pi - t_q)| \ll p^{-1/2}$.

Fact:

$$|\cos^{-1}(1-x) - \cos^{-1}(1-(x+\sqrt{x}))| \ll x^{1/5}.$$

So roughly,

$$|\theta_p - \theta_q| \ll p^{-1/5},$$

After taking $\cos^{-1}$, all we can guarantee is that

$$|\theta_p - \theta_q| \ll$$

— — — — — — — — — — — — — —

Let's think systematically. We're picking t_1 and t_2 close to 1, which is where $(\cos^{-1})'$ blows up. But there shouldn't be very many of them close to 1. Aka,

$$\left| \frac{\#\{p \leqslant x : \theta_p \in [0,t)\}}{\pi(x)} - \int_0^t \mathrm{dST} \right| \ll \frac{1}{\log x}$$

$$\frac{\#\{p \leqslant x : \theta_p \in [0,t)\}}{\pi(x)} \ll t^2 + \frac{1}{\log x}.$$

We want to know, given x, how small the smallest $\theta_p, p \leqslant x$ is. Roughly, for what t is

$$\#\{p \leqslant x : \theta_p \in [0,t)\} < 1?$$

We already know that

$$\#\{p \leqslant x : \theta_p \in [0,t)\} \ll \frac{x}{\log x}\left(t^2 + \frac{1}{\log x}\right).$$

This is frustrating, because it means, essentially, that our convergence to the Sato–Tate measure is so slow (by design) that we can't *ever* guarantee that no θ_p lies in some small interval. But there's something easier. For each $p \leqslant x$, we start by choosing $a_p \in \mathbf{Z}$. How close can a_p be to $2\sqrt{p}$? Numerical experiments **(prove this!)** show that for $t_p = \frac{a_p}{2\sqrt{p}}$, we have

$$|1 - t_p| \gg p^{-1/2}.$$

This is key! That means θ_p won't be too small. In particular, we can control how close θ_p and θ_q will be.

We already have chosen θ_p. We want to choose a_q so that $\cos^{-1}(\frac{a_q}{2\sqrt{q}}) \approx \pi - \theta_p$, i.e.

$$\frac{a_q}{2\sqrt{q}} \approx \sin(\theta_p).$$

We can ensure

$$\left| \frac{a_q}{2\sqrt{q}} - \cos(\pi - \theta_p) \right| \ll p^{-1/2}.$$

Moreover, we know that $|\pm 1 - \frac{a_q}{2\sqrt{q}}| \gg q^{-1/2}$, and likewise for a_p. Thus,

$$|\theta_p - \theta_q| = \left| \cos^{-1}\left(\frac{a_p}{2\sqrt{p}}\right) - \pi + \cos^{-1}\left(\frac{a_q}{2\sqrt{q}}\right) \right| \ll p^{-1/2} \cdot ?$$

Good news: numerical experiments show that we can get very good approximation to $U_1(\theta_q) \approx -U_1(\theta_p)$ for $p < q$ successive primes. This is fantastic!

Numerical experiments suggest that we can enforce

$$|U_1(\theta_p) + U_1(\theta_q)| \ll \frac{\log p}{p}.$$

$\square$

Let (X, μ) be a topological measure space. Suppose g is a non-trivial automorphism of X, such that $g_* \mu = \mu$. Suppose $g^2 = 1$. If we want to minimize

$$\left| \sum_{p \leqslant x} f(x_p) \right|,$$

while letting the discrepancy of $\{x_p\}$ vary arbitrarily. Suppose we can find a "good" subset $U \subset X$ such that $X = U \sqcup gU$. Choose $x_{p_{\text{odd}}} \in U$ to control the discrepancy, and then choose $x_{p_{\text{even}}} \approx g(x_{p_{\text{odd}}})$. For any $f \in C^\infty(X)$ such that $g^* f = -f$. Then

$$\sum_{p \leqslant x} f(x_p) = \sum (f(x_{p_{\text{even}}}) + f(x_{p_{\text{odd}}})) \approx \sum 0.$$

We know that near $\theta = 0$,

$$U_n(\theta) = n + C_n \theta^2 + O(\theta^3).$$

(I think this will hold for any f with $\int f = 0$ and $f(\pi - \theta) = f(\theta)$.)

3 Precise method

Let $\{p_1, p_2, \ldots, \}$ be an enumeration of the rational primes. Given $x \in \mathbf{R}$, write $\sum_{p_{\text{odd}} \leqslant x} a_p$ for the sum of all a_p for $p_i \leqslant x$ with i odd, and similarly for $\sum_{p_{\text{even}} \leqslant x}$. Suppose we have chosen $\theta_{p_{\text{odd}}} \in [0, \pi/2)$ so that $\operatorname{disc}(\{\theta_{p_{\text{odd}}}\}_{p_{\text{odd}} \leqslant x})$ decays as desired. Suppose we choose $\theta_{p_{\text{even}}} \approx \pi - \theta_{p_{\text{odd}}}$. That is, for $p < q$ successive primes with $p = p_i$, i odd, we'll choose $\theta_q \approx \pi - \theta_p$.

We know that $\theta_p = \cos^{-1}\left(\frac{a_p}{2\sqrt{p}}\right)$ for some $a_p \in \mathbf{Z}$ with $|a_p| \leqslant 2\sqrt{p}$. We want to choose $\theta_q \approx \pi - \theta_p$, i.e.

$$\cos^{-1}\left(\frac{a_q}{2\sqrt{q}}\right) \approx \pi - \cos^{-1}\left(\frac{a_p}{2\sqrt{p}}\right)$$
$$\frac{a_q}{2\sqrt{q}} \approx -\frac{a_p}{2\sqrt{p}}.$$

since $\cos(\pi - \cos^{-1}(x)) = -x$. We can guarantee that

$$\left| \frac{a_q}{2\sqrt{q}} + \frac{a_p}{2\sqrt{p}} \right| \leqslant \frac{1}{\sqrt{q}}.$$

Claim: if x, y are "further than ϵ" from ± 1 and $|x-y| < \epsilon$, then $|\cos^{-1}(x) - \cos^{-1}(y)| \leqslant \sqrt{\epsilon}$. (Have checked with Wolfram Alpha, prove later.)

In conclusion, for each successive primes $p = p_{\text{odd}} < q = p_{\text{even}}$, if there is $\theta_p \in \Theta_p(l)$ chosen already, we can also choose $\theta_q \in \Theta_q(l)$ so that

$$|\theta_q - (\pi - \theta_p)| \ll l p^{-1/4}.$$

This is all that is needed, since we're looking at f that is of the form

$$f(\theta) = f(0) + C\theta^2 + O(\theta^3)$$

for θ close to zero. (In fact, this is true for *all* smooth, Weyl-invariant f, whether or not they satisfy $f(\theta) = -f(\pi - \theta)$.) The squaring "pushes the difference" back to $p^{-1/2}$. That is, for θ, φ close to zero, but at least ϵ away from zero, we have

$$|f(\theta) - f(\varphi)| \ll |\theta - \varphi|^2.$$

Now the question is, if $\theta_q \approx \pi - \theta_p$, how close is the discrepancy of $\{\theta_{p_{\text{odd}}}\}$ and $\{\theta_{p_{\text{even}}}\}$?

Better, how close are

$$\#\{p_{\text{odd}} \leqslant x : \theta_{p_{\text{odd}}} \leqslant t\} \qquad \text{and} \qquad \#\{p_{\text{odd}} \leqslant x : \theta_{p_{\text{odd}}} \leqslant t\}?$$

We know that $|\theta_p - \theta_q| \ll p^{-1/4}$. Actually, all we need is that if $\mathrm{disc}(\{\theta_{p_{\text{odd}}}\}) \to 0$, then also $\mathrm{disc}(\{\theta_{p_{\text{even}}}\}) \to 0$.

Suppose we have two sequences $\{x_n\}$ and $\{y_n\}$ such that $\mathrm{disc}(\{x_n\}_{n \leqslant N}) \sim \frac{1}{\log N}$, and also $|x_n - y_n| \leqslant n^{-1/4}$. For some really big N, choose $M < N$, ideally $M \approx \log N$.

Look at

$$\limsup_{N \to \infty} \mathrm{disc}(\{y_n\}_{M \leqslant n \leqslant N}) \leqslant M^{-1/4}.$$

With complete generality, we have:

$$|\mathrm{disc}(\{x_n\}_{n \leqslant N}) - \mathrm{disc}(\{x_n\}_{M \leqslant n \leqslant N})| \ll \frac{1}{M}$$

This is all we need.

— — — — — — — — — — — —

Theorem 2. *Let $\boldsymbol{x}$ and $\boldsymbol{y}$ be sequences in $\mathbf{R}$ such that*

1. $\mathrm{disc}(\boldsymbol{x}^N, \nu) \to 0$.
2. $\|\boldsymbol{x}_{\geqslant N} - \boldsymbol{y}_{\geqslant N}\|_\infty \to 0$.

Then $\mathrm{disc}(\boldsymbol{y}^N, \nu) \to 0$ *as well.*

Proof. Recall our notation. First, $\boldsymbol{x}^N$ is the measure

$$\int f \, \mathrm{d}\boldsymbol{x}^N = \frac{1}{N} \sum_{n \leqslant N} f(x_n),$$

while $\boldsymbol{x}_{\geqslant N} = (x_N, x_{N+1}, \dots)$. The discrepancy between two measures μ and ν on $\mathbf{R}$ is

$$\operatorname{disc}(\mu,\nu) = \sup_{x\in\mathbf{R}} |\operatorname{cdf}_\mu(x) - \operatorname{cdf}_\nu(x)|.$$

Finally, the norm in line two is the supremum norm

$$\|x_{\geqslant N} - y_{\geqslant N}\|_\infty = \sup_{n\geqslant N} |x_n - y_n|.$$

Note that the discrepancy satisfies the triangle inequality, namely

$$\operatorname{disc}(\mu,\xi) \leqslant \operatorname{disc}(\mu,\nu) + \operatorname{disc}(\nu,\xi).$$

Moreover, for $M < N$, write $\boldsymbol{x}^{M,N}$ for the measure

$$\int f \,\mathrm{d}\boldsymbol{x}^{M,N} = \frac{1}{N-M} \sum_{M<n\leqslant N} f(x_n).$$

First, note that $\operatorname{disc}(\boldsymbol{x}^N, \boldsymbol{x}^{M,N}) \leqslant \frac{M}{N}$, because the two finite sets in question differ by M elements, so the cumulative distribution functions can differ by at most $\frac{M}{N}$. Now note that:

$$\operatorname{disc}(\boldsymbol{y}^N,\nu) \leqslant \operatorname{disc}(\boldsymbol{y}^N,\boldsymbol{y}^{M,N}) + \operatorname{disc}(\boldsymbol{y}^{M,N},\boldsymbol{x}^{M,N}) + \operatorname{disc}(\boldsymbol{x}^{M,N},\boldsymbol{x}^N) + \operatorname{disc}(\boldsymbol{x}^N,\nu).$$

Let $M = o(N)$ but still have $M \to \infty$. Then the first and third terms converge to zero. The fourth term converges to zero by hypothesis, so all we need is to consider the second term.

[Proof is invalid in this generality! We only have: if ν is continuous measure, then

$$|\operatorname{disc}(\boldsymbol{x}^N,\nu) - \operatorname{disc}(\boldsymbol{y}^N,\nu)| \ll_\nu \|\boldsymbol{x}_{\leqslant N} - \boldsymbol{y}_{\leqslant N}\|_\infty.$$

] $\square$

More generally,

$$\operatorname{disc}(\mu_{S\sqcup T},\mu_S) = \|\operatorname{cdf}_{\mu_{S\sqcup T}} - \operatorname{cdf}_{\mu_S}\|_\infty \leqslant \frac{\#T}{\#S}.$$

Moreover,

$$\operatorname{disc}(\boldsymbol{y}^N,\nu) \leqslant \operatorname{disc}(\boldsymbol{y}^N,\boldsymbol{x}^N) + \operatorname{disc}(\boldsymbol{x}^N,\nu).$$

So to prove $\operatorname{disc}(\boldsymbol{y}^N,\nu) \to 0$, it suffices in fact to prove that $\operatorname{disc}(\boldsymbol{x}^N,\boldsymbol{y}^N) \to 0$. Here, we have a chain of inequalities (where we write $\boldsymbol{x}^{M,N}$ for the measure corresponding to $\{x_n\}_{M\leqslant N\leqslant M}$:

$$\operatorname{disc}(\boldsymbol{x}^N,\boldsymbol{y}^N) \leqslant \operatorname{disc}(\boldsymbol{x}^N,\boldsymbol{x}^{M,N}) + \operatorname{disc}(\boldsymbol{x}^{M,N},\boldsymbol{y}^{M,N}) + \operatorname{disc}(\boldsymbol{y}^{M,N},\boldsymbol{y}^N).$$

Now, it is known that

$$\operatorname{disc}(\boldsymbol{x}^{M,N},\boldsymbol{y}^{M,N}) \ll M^{-1/4},$$

Now, $\mathrm{disc}(\boldsymbol{x}^N, \boldsymbol{x}^{M,N}) \leqslant \frac{M}{N-M}$, so we can set $M \approx \sqrt{N}$ to get

$$\mathrm{disc}(\boldsymbol{x}^N, \boldsymbol{y}^N) \ll \frac{\sqrt{N}}{N-\sqrt{N}} + N^{-1/4} \to 0.$$

— — — — — — — — — — — — — — — —

$$|\,\mathrm{disc}(\boldsymbol{y}^{M,N}, \nu) - \mathrm{disc}(\boldsymbol{x}^{M,N}, \nu)| \ll_\nu \|\boldsymbol{x}_{\geqslant M} - \boldsymbol{y}_{\geqslant M}\|_\infty.$$

Thus

$$\mathrm{disc}(\boldsymbol{y}^{M,N}, \nu) \ll_\nu \mathrm{disc}(\boldsymbol{x}^{M,N}, \nu) + \|\boldsymbol{x}_{\geqslant M} - \boldsymbol{y}_{\geqslant M}\|_\infty$$

Lemma 1. *Let $\boldsymbol{x}$ be a sequence in $\mathbf{R}$. For $M < N$,*

$$\mathrm{disc}(\boldsymbol{x}^N, \boldsymbol{x}^{M,N}) \leqslant \frac{2M}{N}.$$

Proof. We know that

$$|\#\{n \leqslant N : x_n \leqslant t\} - \#\{M < n \leqslant N : x_n \leqslant t\}| \leqslant M.$$

Thus for all $t \in \mathbf{R}$,

$$|\,\mathrm{cdf}_{\boldsymbol{x}^N}(t) - \frac{N-M}{N}\,\mathrm{cdf}_{\boldsymbol{x}^{M,N}}(t)| \leqslant \frac{M}{N}.$$

Now just do a simple computation, with $|\cdot| = \|\cdot\|_\infty$:

$$\begin{aligned} |\,\mathrm{cdf}_{\boldsymbol{x}^N} - \mathrm{cdf}_{\boldsymbol{x}^{M,N}}\,| &\leqslant \left|\mathrm{cdf}_{\boldsymbol{x}^N} - \frac{N-M}{N}\,\mathrm{cdf}_{\boldsymbol{x}^{M,N}}\right| + \left|\frac{N-M}{N}\,\mathrm{cdf}_{\boldsymbol{x}^{M,N}} - \mathrm{cdf}_{\boldsymbol{x}^{M,N}}\right| \\ &\leqslant \frac{M}{N} + \frac{M}{N}. \end{aligned}$$

□

Lemma 2. *Let $\boldsymbol{x}$ and $\boldsymbol{y}$ be sequences in $\mathbf{R}$. Suppose $\nu = f\,\mathrm{d}x$ for a continuous function f. Then*

$$|\,\mathrm{disc}(\boldsymbol{x}^N, \nu) - \mathrm{disc}(\boldsymbol{y}^N, \nu)| \leqslant 2\epsilon\|f\|_\infty + \frac{\#\{n \leqslant N : |x_n - y_n| > \epsilon\}}{N}.$$

Proof. It is actually sufficient to prove that

$$\mathrm{disc}(\boldsymbol{x}^N, \nu) \leqslant \mathrm{disc}(\boldsymbol{y}^N, \nu) + 2\epsilon\|f\|_\infty + \frac{\#\{n \leqslant N : |x_n - y_n| > \epsilon\}}{N}.$$

□