Analysis, Random Walks and Groups

Exercise sheet 1 solutions

These model solutions are the same models used by Tuomas in Manchester, but some exercises are omitted, attached into another exercises or separeted into separate exercises. I (Kai) might have commented somewhere with red color if I think it is in place. Corrections and improvements are welcome.

1. (5pts.)

In the **weak Borel suffle** one lifts the top card from a deck of 52 cards and inserts it into the deck in a random position.

(a) For $0 \leq j \leq 51$ determine the permutation $\sigma_j \in S_52$ corresponding to the outcome of placing the top card in the jth position amongst the remaining cards. (The 0th position is on top of the remaining cards and the 51st position is on the bottom of the remaining cards.)

(b) Supposing that one performs consecutive weak Borel shuffles with $0 \leq j \leq 51$ chosen uniformly at random (e.g. by rolling a 52-sided die each time) what is the probability that the card on the top of the deck before shuffling is on top of the deck after two shuffles?

Solution 1.a

We need to move the first card (at slot 0) to a random slot from $0, 1, \ldots, 51$. For $j = 0, 1, \ldots, 51$ with probability $1/52$ and preserve the order. For this purpose, first define a permutation

$$\sigma_k(j) = \begin{cases} k & \text{if } j = 0; \\ j-1 & \text{if } 1 \leq j \leq k; \\ j & \text{if } k < j \leq 51; \end{cases}$$

See Figure 1 for an illustration of σ_k.

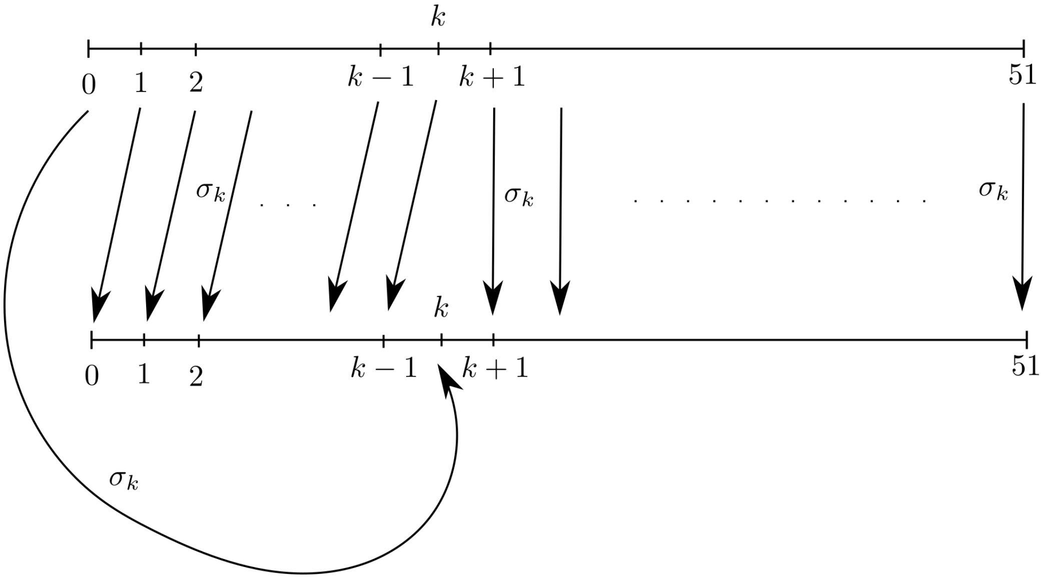

FIGURE 1. Permutation σ_k moves the 0:th card to the k:th slot and adjusts the other cards to keep the order.

To solve the question, we can then define the random permutation $\sigma \in S_{52}$ by choosing $k \in \{0, 1, \dots, \cancel{52}\ 51\}$ uniformly with probability $1/52$ and setting $\sigma = \sigma_k$.

Solution 1.b

We are asking the probability of the event that after choosing $k_1 \in \{0, 1, \dots, \cancel{52}\ 51\}$ with probability $1/52$ and then $k_2 \in \{0, 1, \dots, \cancel{52}\ 51\}$ with probability $1/52$, independently of each other, we end up to a random product permutation

$$\sigma_{k_2}\sigma_{k_1}$$

that satisfies

$$\sigma_{k_2}(\sigma_{k_1}(0)) = 0.$$

Let us first look at when, if $k \in \{0, 1, \dots, 51\}$ is given, then for which j we can have $\sigma_k(j) = 0$:

(a) If $k = 0$, then the only possibility is $j = 0$, as then $\sigma_0(0) = 0$.
(b) If $k \geq 1$, then the only possibility is $j = 1$, as then $\sigma_k(1) = 0$.

Thus by (a) and (b) in order for the event

$$\sigma_{k_2}(\sigma_{k_1}(0)) = 0$$

to occur,

(1) either $k_2 = 0$ and $\sigma_{k_1}(0) = 0$, that is, $k_1 = 0$;
(2) or $k_2 \geq 1$ and $\sigma_{k_1}(0) = 1$, that is, $k_1 = 1$.

Hence the event

$$\sigma_{k_2}(\sigma_{k_1}(0)) = 0$$

occurs if and only if either $k_1 = k_2 = 0$ or $[k_1 = 1 \text{ and } k_2 \geq 1]$.

Since we choose k_1 and k_2 with probabilities $1/52$, independently of each other, this means that the event $k_1 = k_2 = 0$ happens with probability

$$\frac{1}{52^2},$$

and the event $[k_1 = 1 \text{ and } k_2 \geq 1]$ happens with probability

$$\frac{1}{52} \cdot \frac{51}{52} = \frac{51}{52^2}.$$

Therefore, the answer to the question is

$$\frac{1}{52^2} + \frac{51}{52^2} = \frac{52}{52^2} = \frac{1}{52}.$$

2. (5pts)

Prove the variational formula for the total variation distance between two probability distributions μ, ν in $\mathbb{Z}_p$:

$$d(\mu,\nu) = \frac{1}{2} \max\{|\mu(f) - \nu(f)| : \|f\|_\infty \leq 1, f : \mathbb{Z}_p \to \mathbb{R}\}.$$

Hint: For the more difficult upper bound, first define a suitable function $g : \mathbb{Z}_p \to \mathbb{R}$ using the set $B = \{t \in \mathbb{Z}_p : \mu(t) \geq \nu(t)\}$ such that you can ensure $\mu(g) - \nu(g) = \sum_{t \in \mathbb{Z}_p} |\mu(t) - \nu(t)|$ (try to think something quite simple for g), and then use the $L1$ identity from the lecture notes.

Solution 2.

Let us first prove

$$d(\mu,\nu) \leq \frac{1}{2} \max\{|\mu(f) - \nu(f)| : \|f\|_\infty \leq 1, f : \mathbb{Z}_p \to \mathbb{R}\}.$$

To do this, we will construct a function g such that $\frac{1}{2}|\mu(g) - \nu(g)|$ realises the $d(\mu,\nu)$. Write

$$B = \{t \in \mathbb{Z}_p : \mu(t) \geq \nu(t)\}.$$

Define a function

$$g(t) = \begin{cases} +1, & t \in B \\ -1, & t \notin B. \end{cases}$$

Then we have

$$\begin{aligned} \mu(g) - \nu(g) &= \sum_{t \in \mathbb{Z}_p} g(t)\mu(t) - \sum_{t \in \mathbb{Z}_p} g(t)\nu(t) \\ &= \sum_{t \in \mathbb{Z}_p} g(t)(\mu(t) - \nu(t)) \\ &= \sum_{t \in B} g(t)(\mu(t) - \nu(t)) + \sum_{t \notin B} g(t)(\mu(t) - \nu(t)) \end{aligned}$$

Here if $t \in B$ we have by the definition of B that

$$g(t)(\mu(t) - \nu(t)) = \mu(t) - \nu(t) \geq 0$$

and if $t \notin B$ we have by the definition of B that

$$g(t)(\mu(t) - \nu(t)) = -(\mu(t) - \nu(t)) \geq 0.$$

Hence

$$\sum_{t \in B} g(t)(\mu(t)-\nu(t)) + \sum_{t \notin B} g(t)(\mu(t)-\nu(t)) = \sum_{t \in B} |\mu(t)-\nu(t)| + \sum_{t \notin B} |\mu(t)-\nu(t)| = \sum_{t \in \mathbb{Z}_p} |\mu(t)-\nu(t)|.$$

This in particular proves that

$$\mu(g) - \nu(g) = \sum_{t \in \mathbb{Z}_p} |\mu(t) - \nu(t)|,$$

making $\mu(g) - \nu(g) \geq 0$. Thus by the L^1 formula for the total variation distance we obtain

$$d(\mu, \nu) = \frac{1}{2} \sum_{t \in \mathbb{Z}_p} |\mu(t) - \nu(t)| = \frac{1}{2} |\mu(g) - \nu(g)|.$$

Since

$$\|g\|_\infty = \max\{|g(t) : t \in \mathbb{Z}_p\} = 1,$$

we have proved

$$d(\mu, \nu) \leq \frac{1}{2} \max\{|\mu(f) - \nu(f)| : \|f\|_\infty \leq 1, f : \mathbb{Z}_p \to \mathbb{R}\}.$$

We still need to prove the reverse direction:

$$d(\mu, \nu) \geq \frac{1}{2} \max\{|\mu(f) - \nu(f)| : \|f\|_\infty \leq 1, f : \mathbb{Z}_p \to \mathbb{R}\}.$$

To do this, fix any $f : \mathbb{Z}_p \to \mathbb{R}$ with $\|f\|_\infty \leq 1$. Then $f(t) \leq 1$ for all $t \in \mathbb{Z}_p$. Thus we have by the triangle inequality

$$\begin{aligned}
|\mu(f) - \nu(f)| &= \Big| \sum_{t \in \mathbb{Z}_p} f(t)\mu(t) - \sum_{t \in \mathbb{Z}_p} f(t)\nu(t) \Big| \\
&= \Big| \sum_{t \in \mathbb{Z}_p} f(t)(\mu(t) - \nu(t)) \Big| \\
&\leq \sum_{t \in \mathbb{Z}_p} |f(t)||\mu(t) - \nu(t)| \\
&\leq \sum_{t \in \mathbb{Z}_p} 1 \cdot |\mu(t) - \nu(t)| \\
&= 2d(\mu, \nu),
\end{aligned}$$

where in the last line we used the L^1 identity for the total variation distance. Hence the claim follows as $f : \mathbb{Z}_p \to \mathbb{R}$ is arbitrary.

Let $0 < \alpha < 1$, integer $p \geq 2$ and define the following probability distribution on $\mathbb{Z}_p$:

$$\mu_\alpha = \alpha\delta_1 + (1 - \alpha)\delta_{-1}.$$

3.

Find the probabilities of the events:

(a) *"a randomly chosen $t \in \mathbb{Z}_p$ with respect to μ_α is even"*

(b) *"a randomly chosen $t \in \mathbb{Z}_p$ with respect to μ_α is prime"*

Solution 3.a

For the first event, define the set

$$A = \{t \in \mathbb{Z}_p : t \text{ is even}\}.$$

Then we are asking the measure $\mu_\alpha(A)$. The answer will depend on whether p is even or odd.

Note that $-1 = p-1$. Thus:

- If p is odd, then $p-1$ is even, that is, $\delta_{p-1}(A) = 1$, so

$$\mu_\alpha(A) = \alpha \cdot 0 + (1-\alpha) \cdot 1 = 1-\alpha.$$

Hence if p is odd, then the answer is a randomly chosen $t \in \mathbb{Z}_p$ with respect to μ_α is a even with probability $1-\alpha$.

- If p is even, then $p-1$ is odd, that is, $\delta_{p-1}(A) = 0$, so

$$\mu_\alpha(A) = \alpha \cdot 0 + (1-\alpha) \cdot 0 = 0.$$

Hence if p is even, then the answer is a randomly chosen $t \in \mathbb{Z}_p$ with respect to μ_α is a even with probability 0.

Solution 3.b

¸ Now for the second event, define the set

$$B = \{t \in \mathbb{Z}_p : t \text{ is prime}\}.$$

We need to find the measure $\mu_\alpha(B)$.

Note that primes are always strictly greater than 1, so $1 \notin B$.

- If $p-1$ is a prime, then $\delta_{p-1}(B) = 1$ so

$$\mu_\alpha(B) = \alpha \cdot 0 + (1-\alpha) \cdot 1 = 1-\alpha.$$

Hence if $p-1$ is a prime, then the answer is a randomly chosen $t \in \mathbb{Z}_p$ with respect to μ_α is a prime with probability $1-\alpha$.

- If $p-1$ is not a prime, then $\delta_{p-1}(B) = 1$ so

$$\mu_\alpha(B) = \alpha \cdot 0 + (1-\alpha) \cdot 0 = 0.$$

Hence if $p-1$ is not a prime, then the answer is a randomly chosen $t \in \mathbb{Z}_p$ with respect to μ_α is a prime with probability 0.

4.

Define a function $f : \mathbb{Z}_p \to \mathbb{C}$ by

$$f(t) = \begin{cases} +1; & t \text{ is even;} \\ -1; & t \text{ is odd.} \end{cases}$$

Find the integral (i.e. expectation) $\mu_\alpha(f)$ of f.

Solution 4.

¸Firstly we know that $-1 = p-1$. Thus the value of $\mu_\alpha(p-1)$ depends whether p is even or odd.

Case 1: p is odd. Then $p-1$ is even so

$$\mu_\alpha(f) = \sum_{t\in\mathbb{Z}_p} f(t)\mu_\alpha(t) = -1\cdot\alpha + 1\cdot(1-\alpha) = \cancel{2-\alpha}\ 1-2\alpha.$$

Case 2: p is even. Then $p-1$ is odd so

$$\mu_\alpha(f) = \sum_{t\in\mathbb{Z}_p} f(t)\mu_\alpha(t) = -1\cdot\alpha + (-1)\cdot(1-\alpha) = -1.$$

The number $\mu_\alpha(f)$ tells us the average sign of a random number $t\in\mathbb{Z}_p$ (where sign is $+1$ if even and -1 if odd).

5.

Define a function $f:\mathbb{Z}_p\to\mathbb{C}$ by

$$f(t) = \begin{cases} +1; & t \text{ is even;} \\ -1; & t \text{ is odd.} \end{cases}$$

Find the integral (i.e. expectation) $\lambda(f)$ of f with respect to the uniform measure.

Solution 5.

If p is even, we did this already in the lectures, so there are $p/2$ even and $p/2$ odd numbers in $\{0,1,\dots,p-1\}$. Thus

$$\lambda(f) = \sum_{t\in\mathbb{Z}_p} f(t)\lambda(t) = \frac{p}{2}\cdot\frac{1}{p} - \frac{p}{2}\cdot\frac{1}{p} = 0.$$

If p is odd, then $p-1$ is even (and we define 0 to be even), so there are in total $(p-1)/2$ odd numbers and $(p-1)/2+1$ even numbers. Hence

$$\lambda(f) = \sum_{t\in\mathbb{Z}_p} f(t)\lambda(t) = \Big(\frac{p-1}{2}+1\Big)\cdot\frac{1}{p} - \frac{p-1}{2}\cdot\frac{1}{p} = \frac{1}{p}.$$

6.

The **distance** between two points $t,s\in\mathbb{Z}_p$ is

$$dist(t,s) = \min\{t\oplus(-s), (-t)\oplus s\},$$

which measures the shortest distance between t and s at the dinner table. A function $f:\mathbb{Z}_p\to\mathbb{R}$ is **Lipschitz** if there exists $L\geq 0$ (it exists for any f, see below), is denoted as $\mathrm{Lip}(f)$. The **earth mover's distance** or also known as the first **Wasserstein distance** $W_1(\mu,\nu)$ of two probability distributions μ,ν in $\mathbb{Z}_p$ is given by

$$W_1(\mu,\nu) = \max\Big\{|\mu(f)-\nu(f)| : \mathrm{Lip}(f)\leq 1, f:\mathbb{Z}_p\to\mathbb{R}\Big\}.$$

Wasserstein distance appears commonly in a field called **mass transportation theory**, which has many applications throughout economics, physics and mathematics of PDEs. We can state it in our course's language as follows. Given a probability distribution μ on $\mathbb{Z}_p$ and a map $T : \mathbb{Z}_p \to \mathbb{Z}_p$, define the **push-forward** $T_*\mu$ by the formula

$$T_*\mu(t) = \mu(T^{-1}\{t\}), \quad t \in \mathbb{Z}_p,$$

where $T^{-1}\{t\} = \{s \in \mathbb{Z}_p : T(t) = s\}$ is the pre-image of the singleton $\{t\}$. Given two probability distributions μ, ν in $\mathbb{Z}_p$, a mapping $T : \mathbb{Z}_p \to \mathbb{Z}_p$ that maps μ onto ν, $T_*\mu = \nu$, is called an **optimal transportation** if it minimises the "cost":

$$\int \mathrm{dist}(t, T(t))\, d\mu(t). \qquad \Big(= \sum_{t \in \mathbb{Z}_p} \mathrm{dist}(t, T(t))\, \mu(t) \Big)$$

The **Monge-Kantorovich duality theorem** says the minimal cost is the first Wasserstein distance:

$$\min \Big\{ \int \mathrm{dist}(t, T(t))\, d\mu(t) : T_*\mu = \nu \Big\} = W_1(\mu, \nu),$$

the proof can be found from literature on optimal transportation theory.

(a) Prove that any $f : \mathbb{Z}_p \to \mathbb{R}$ is Lipschitz. Which functions $f : \mathbb{Z}_p \to \mathbb{R}$ satisfy $\mathrm{Lip}(f) = 0$?

(b) Prove that

$$d(\mu, \nu) \leq W_1(\mu, \nu),$$

where $d(\mu_{,}, \nu)$ is the total variation distance.

(c) Fix $s \in \mathbb{Z}_p$. Define a transportation map $T : \mathbb{Z}_p \to \mathbb{Z}_p$ by $T(t) = s,\ t \in \mathbb{Z}_p$. Verify that the uniform distribution $\lambda(t) = \frac{1}{p},\ t \in \mathbb{Z}_p$, satisfies $T_*\lambda = \delta_s$, where δ_s is the singular distribution at s. What is the cost of transporting the uniform mass λ to a point δ_s? That is, find the cost

$$\int_{\mathbb{Z}_p} dist(t, T(t)) d\lambda(t).$$

Can you transport δ_s to λ?

Solution 6.a

The number

$$0 \leq L := \max_{a \neq b} \frac{|f(a) - f(b)|}{d(a, b)} < \infty$$

as $\mathbb{Z}_p$ is finite since $d(a, b) \geq 1$ when $a \neq b$ and the maximum is over a finite set $\mathbb{Z}_p$. Then

$$|f(t) - f(s)| = \frac{|f(t) - f(s)|}{d(t, s)} \cdot d(t, s) \leq L d(t, s).$$

Moreover, if $t = s$, then as $d(t, s) \geq 0$ we have $|f(t) - f(s)| = 0 \leq L d(t, s)$, so f is Lipschitz.

After a moment, we notice that the constant functions cause $L = 0$.

Solution 6.b

Take $f : \mathbb{Z}_p \to \mathbb{R}$ such that $\|f\|_\infty \leq 1$. To prove the claim, it will be enough if we construct a Lipschitz function $g : \mathbb{Z}_p \to \mathbb{R}$ with $\mathrm{Lip}(g) \leq 1$ such that

$$|\mu(f) - \nu(f)| \leq 2|\mu(g) - \nu(g)|.$$

For this purpose, let us first define a number

$$L := \max_{t \neq s} \frac{|f(t) - f(s)|}{d(t, s)}.$$

Since for $t \neq s$ the distance $1 \leq d(t,s)$ and $|f(t) - f(s)| \leq 2\|f\|_\infty \leq 2$ we have

$$0 \leq L \leq 2.$$

Define a function $g : \mathbb{Z}_p \to \mathbb{R}$ by

$$g(t) = f(t)/L, \quad t \in \mathbb{Z}_p.$$

Then for $t \neq s$ we have by the definition of L that

$$|g(t) - g(s)| = \frac{1}{L}|f(t) - f(s)| = \frac{1}{L} \cdot \frac{|f(t) - f(s)|}{d(t,s)} \cdot d(t,s) \leq d(t,s).$$

Moreover, for $t = s$ we have

$$|g(t) - g(s)| = 0 \leq d(t,s).$$

Thus g is Lipschitz with $\mathrm{Lip}(g) \leq 1$. Lastly, since

$$\mu(g) = \sum_{t \in \mathbb{Z}_p} g(t)\mu(t) = \sum_{t \in \mathbb{Z}_p} \frac{f(t)}{L}\mu(t) = L^{-1}\mu(f),$$

and similarly for $\nu(g)$, we see that

$$|\mu(f) - \nu(f)| = L|L^{-1}\mu(f) - L^{-1}\nu(f)| = L|\mu(g) - \nu(g)| \leq 2|\mu(g) - \nu(g)|,$$

which gives the claim.

Notice that in the previous proof, L can be zero. Hence when defining g, where can be a division by zero. This is the case when the function is constant. This should be handled separately, but this should not be too hard, just choose $g = f$.´

Solution 6.c

Fix $t \in \mathbb{Z}_p$. We have

$$\lambda(T^{-1}\{t\}) = \lambda(\varnothing) = 0;$$

if $t \neq s$ and

$$\lambda(T^{-1}\{t\}) = \lambda(\mathbb{Z}_p) = 1$$

if $t = s$. Therefore, $\lambda(T^{-1}\{t\}) = \delta_s(t)$, for all $t \in \mathbb{Z}_p$, as we claimed.

By the definition of the integral

$$\int \mathrm{dist}(t, T(t))\, d\lambda(t) = \sum_{t \in \mathbb{Z}_p} \mathrm{dist}(t, T(t))\lambda(t) = \sum_{t \in \mathbb{Z}_p} \mathrm{dist}(t,s)\frac{1}{p} = \frac{1}{p}\sum_{t \in \mathbb{Z}_p} \mathrm{dist}(t,s) = *$$

If p is even, then we see that

$$\sum_{t \in \mathbb{Z}_p} \mathrm{dist}(t,s) = \sum_{k=1}^{p/2} 1 + \sum_{k=1}^{p/2} 1 = \frac{p(p+1)}{2}$$

by taking the both directions from s to either clockwise or counterclockwise. Hence the cost is

$$\int \mathrm{dist}(t, T(t))\, d\lambda(t) = \frac{1}{p}\frac{p(p+1)}{2} = \frac{p+1}{2}.$$

If p is odd, then we have one less distance in the sums, so we have

$$\sum_{t \in \mathbb{Z}_p} \mathrm{dist}(t,s) = \sum_{k=1}^{(p-1)/2} 1 + \sum_{k=1}^{(p-1)/2} 1 = \frac{(p-1)p}{2}.$$

Hence the cost of transporting λ to δ_s is

$$\int \operatorname{dist}(t, T(t))\, d\lambda(t) = \frac{1}{p}\frac{(p-1)p}{2} = \frac{p-1}{2}.$$

I think the previous is wrong. For example, take $p = 2$. Then the cost should be

$$\frac{1}{2} \neq \frac{2+1}{2}$$

. Similarly, if $p = 3$, the cost should be

$$\frac{2}{3} \neq \frac{3-1}{2} = 1$$

. I have something like

IF p is even:

$$\begin{aligned} *= \frac{1}{p}\Big[\sum_{k=1}^{\frac{p}{2}} k + \sum_{k=1}^{\frac{p}{2}-1} k\Big] &= \frac{1}{p}\Big[\frac{p}{2}\frac{\frac{p}{2}+1}{2} + (\frac{p}{2}-1)\frac{\frac{p}{2}-1+1}{2}\Big] \\ &= \frac{1}{p}\Big[\frac{p(\frac{p}{2}+1)}{4} + \frac{p}{2}\frac{p-2}{4}\Big] \\ &= \frac{\frac{p}{2}+1}{4} + \frac{p-2}{8} \\ &= \frac{p+2+p-2}{8} \\ &= \frac{2p}{8} \\ &= \frac{p}{4} \end{aligned}$$

IF p is odd:

$$\begin{aligned} *= \frac{1}{p}\Big[2\sum_{k=1}^{\frac{p-1}{2}} k\Big] &= \frac{2}{p}(\frac{p-1}{2})(\frac{\frac{p-1}{2}+1}{2}) \\ &= \frac{p-1}{p}\frac{p+1}{4} \\ &= \frac{p^2-1}{4p} \end{aligned}$$

It is not possible to transport δ_s to λ since otherwise, if such map $S : \mathbb{Z}_p \to \mathbb{Z}_p$ exists with

$$\lambda = S_* \delta_s.$$

Define $t := S(s) \in \mathbb{Z}_p$. Then $s \in S^{-1}\{t\}$ so we have

$$\delta_s(S^{-1}(t)) = 1.$$

On the other hand, we assumed $\lambda = S_* \delta_s$ so

$$1/p = \lambda(s) = \delta_s(S^{-1}(t)) = 1$$

which is not possible when $p \geq 2$.