Analysis, Random Walks and Groups

Exercise sheet 5: solutions

Homework exercises: Return these for marking to Kai Hippi in the tutorial on Week 6. Contact Kai by email if you cannot return these in-person, and you can arrange an alternative way to return your solutions. Remember to be clear in your solutions, if the solution is unclear and difficult to read, you can lose marks. Also, if you do not know how to solve the exercise, attempt something, you can get awarded partial marks.

Comments etc. by Kai in red color.

1. (5pts)

Let G be a finite group and $\varrho_1 : G \to U(V_{\varrho_1})$ and $\varrho_2 : G \to U(V_{\varrho_2})$ be unitary representations and let $\varphi : V_1 \to V_2$ be a morphism between ϱ_1 to ϱ_2. Prove the following version of Schur's lemma:

(a) if ϱ_1 is irreducible, then φ is either injective or zero;
(b) if ϱ_2 is irreducible, then φ is either surjective or zero.

Solution 1.a

The kernel of φ,

$$\ker(\varphi) = \{v \in V_{\varrho_1} : \varphi(v) = 0\} < V_{\varrho_1}$$

is ϱ_1 invariant: as φ is a morphism, we have for all $v \in \ker(\varphi)$ that

$$\varphi(\varrho_1(x)v) = \varrho_2(x)\varphi(v) = \varrho_2(x)0 = 0$$

so $\varrho_1(x)v \in \ker(\varphi)$ for all $x \in G$. Thus as φ_1 is irreducible, $\ker(\varphi) = \{0\}$, in which case φ is injective as a linear map, or V_{ϱ_1}, in which case it is zero.

Solution 1.b

The range (i.e. image) of φ,

$$\mathrm{im}(\varphi) = \{w \in V_{\varrho_2} : \varphi(v) = w \text{ for some } v \in V_{\varrho_1}\} < V_{\varrho_2}$$

is ϱ_2 invariant. Indeed, fix $w \in \mathrm{im}(\varphi)$ so there exists $v \in V_{\varrho_1}$ such that $\varphi(v) = w$. As φ is a morphism, we thus have

$$\varphi(\varrho_1(x)v) = \varrho_2(x)\varphi(v) = \varrho_2(x)w$$

so $\varrho_2(x)w \in \mathrm{im}(\varphi)$ for all $x \in G$. Thus as φ_2 is irreducible, $\mathrm{im}(\varphi) = \{0\}$, in which case φ is zero, or V_{ϱ_2}, in which case it is surjective by definition.

2. (5pts)

Fix $p \geq 2$ and let H be a subgroup of $\mathbb{Z}_p$. Prove the **Poisson summation formula**: for any $f : \mathbb{Z}_p \to \mathbb{C}$, we have

$$\frac{1}{|H|} \sum_{h \in H} f(h) = \frac{1}{p} \sum_{k \in H^\perp} \widehat{f}(k),$$

where

$$H^{\perp} := \{k \in \mathbb{Z}_p : e^{2\pi i k t/p} = 1 \text{ for all } t \in H\}.$$

Hint: There are couple of ways to do this. One way is to apply the inverse Fourier transform to the function $F(s) = \sum_{h\in H} f(sh)$, and then set $s = 1$, or first verifying the Poisson summation formula for Dirac measures δ_t and then using linearity to extend for all f.

Solution 2.

Let us first verify the formula for Dirac measures and then use these for the asked equality. Let $x \in \mathbb{Z}_p$ and consider $f = \delta_x$. First let us note:

$$\widehat{\delta_x}(k) = e^{-2\pi i x k/p}$$

- **Case:** $H = \mathbb{Z}_p$
  - **Case:** $x \in H$
    Consider
    $$\begin{aligned} H^{\perp} &= \{k \in \mathbb{Z}_p \mid e^{2\pi i k t/p} = 1\ \forall t \in H\} \\ &= \{k \in \mathbb{Z}_p \mid e^{2\pi i k t/p} = 1\ \forall t \in \mathbb{Z}_p\}. \end{aligned}$$
    Clearly $H^{\perp} = \{0\}$ as for $k \in \mathbb{Z}_p, k \neq 0$ it holds that
    $$k\frac{p-1}{p} \notin \mathbb{Z} \iff e^{2\pi i k(p-1)/p} \neq 1$$
    which means that $k \notin H^{\perp}$.

    Clearly
    $$\frac{1}{|H|}\sum_{h\in H} f(h) = \frac{1}{|\mathbb{Z}_p|} = \frac{1}{p}.$$
    And now
    $$\frac{1}{p}\sum_{k\in H^{\perp}} \widehat{f}(k) = \frac{1}{p}\sum_{k\in H^{\perp}} \widehat{\delta_x}(k) = \frac{1}{p}\sum_{k\in H^{\perp}} e^{-2\pi i k x/p} = \frac{1}{p}.$$
    This case is okay.

  - **Case:** $x \notin H$
    Clearly this is not a possible situation as $H = \mathbb{Z}_p$.

- **Case:** $H = \{0\}$
  - **Case:** $x \in H$
    Now $\delta_x = \delta_0$ necessarily. We quickly that
    $$H^{\perp} = \mathbb{Z}_p.$$
    Clearly
    $$\frac{1}{|H|}\sum_{h\in H} f(h) = 1.$$

Also

$$\frac{1}{p}\sum_{k\in H^\perp}\widehat{f}(k)=\frac{1}{p}\sum_{k\in\mathbb{Z}_p}\widehat{\delta}_0(k)=\sum_{k\in H^\perp}1=\frac{1}{p}p=1.$$

This case is okay.

- **Case:** $x\notin H$
  Clearly

$$\frac{1}{|H|}\sum_{h\in H}f(h)=0.$$

Also

$$\frac{1}{p}\sum_{k\in H^\perp}\widehat{f}(k)=\frac{1}{p}\sum_{k\in\mathbb{Z}}e^{-2\pi ikx/p}=0,$$

as $x\neq 0$. This case is okay.

- **Case:** $H\neq\{0\}, H\neq\mathbb{Z}_p$
  - **Case:** $x\in H$

By **thm.2.9.**, if p is prime, then no such H exists we are focusing on here. Hence we may assume p is not a prime number. Also by **thm.2.9.**

$$H=\langle w\rangle$$

for some $w\in\mathbb{Z}_p$ so that $w|p$. From this we see that

$$H=\{aw\mid a\in\{0,...,\frac{p}{w}-1\}\}.$$

We also see that $|H|=\frac{p}{t}$

Consider

$$\begin{aligned}H^\perp&=\{k\in\mathbb{Z}_p\mid e^{2\pi ikt/p=1\ \forall t\in H}\}\\&=\{k\in\mathbb{Z}_p\mid\frac{kt}{p}\ \forall t\in H\}\\&=\{k\in\mathbb{Z}_p\mid k\frac{a}{p/w}\in\mathbb{Z}\ \forall a\in\{0,...,\frac{p}{w}-1\}\}.\end{aligned}$$

Clearly

$$\{r\frac{p}{w}\in\mathbb{Z}_p\mid r\in\mathbb{Z}\}\subseteq H^\perp.$$

Let us assume that there exists $\chi\in H^\perp$ so that $\chi\notin\{r\frac{p}{w}\in\mathbb{Z}_p\mid r\in\mathbb{Z}\}$. Thus

$$\chi\frac{\frac{p}{w}-1}{\frac{p}{w}}\in\mathbb{Z}.$$

Since $\frac{p}{w}$ and $\frac{p}{w}-1$ are coprime integers we have the following result from number theory:

$$\text{if } (\frac{p}{w})|(k(\frac{p}{w}-1)), \text{then } (\frac{p}{w})|k,$$

where k is an integer. This means that

$$\chi=r\frac{p}{w}$$

for some $r \in \mathbb{Z}$. This is a contradiction so

$$H^{\perp} = \{r\frac{p}{w} \in \mathbb{Z}_p \mid r \in \mathbb{Z}\}.$$

Clearly

$$|H^{\perp}| = \frac{p}{\frac{p}{w}} = w.$$

So then:

$$\frac{1}{|H|}\sum_{h\in H} \delta_x(h) = \frac{1}{|H|} = \frac{w}{p}.$$

Now also:

$$\frac{1}{p}\sum_{k\in H^{\perp}} \widehat{f}(k) = \frac{1}{p}\sum_{k\in H^{\perp}} \widehat{\delta_x}(k) = \frac{1}{p}\sum_{k\in H^{\perp}} e^{-2\pi ikx/p} = \frac{1}{p}\sum_{k\in H^{\perp}} 1 = \frac{|H^{\perp}|}{p} = \frac{w}{p}.$$

This case is okay.

– **Case:** $x \notin H$
Clearly again

$$\frac{1}{|H|}\sum_{h\in H} f(h) = 0.$$

Now consider (and use observations regarding $H^{\perp}$ from the last part):

$$\begin{aligned}\frac{1}{p}\sum_{k\in H^{\perp}} \widehat{f}(k) &= \frac{1}{p}\sum_{k\in H^{\perp}} \widehat{\delta_x}(k) \\ &= \frac{1}{p}\sum_{s=0}^{w-1} e^{-2\pi is\frac{p}{w}/p} \\ &= \frac{1}{p}\sum_{s=0}^{w-1} e^{-2\pi is/w} = 0.\end{aligned}$$

This case is okay.

We have confirmed that for $f = \delta_x$, $x \in \mathbb{Z}_p$ the Poisson summation works. Now consider

$$f : \mathbb{Z}_p \to \mathbb{C},\ z \mapsto f(z) := \sum_{a\in\mathbb{Z}_p} c_a\delta_a(z),$$

where $c_a \in \mathbb{C}$. Consider

$$\begin{aligned}
\frac{1}{|H|}\sum_{h\in H} f(h) &= \frac{1}{|H|}\sum_{h\in H}\sum_{b\in\mathbb{Z}_p} c_b\delta_b(h) \\
&= \sum_{b\in\mathbb{Z}_p} c_b \frac{1}{|H|}\sum_{h\in H}\delta_b(h) \\
&= \sum_{b\in\mathbb{Z}_p} c_b \frac{1}{p}\sum_{k\in H^\perp}\widehat{\delta_b}(k) \\
&= \frac{1}{p}\sum_{k\in H^\perp}\widehat{\sum_{b\in\mathbb{Z}_p} c_b\delta_b}(k) \\
&= \frac{1}{p}\sum_{k\in H^\perp}\widehat{f}(k).
\end{aligned}$$

We have confirmed the Poisson summation formula.

Further exercises: Attempt these before the tutorial, they are not marked and will be discussed in the tutorial. If you cannot attend the tutorial, but want to do the attendance marks, you can return your attempts to these before the tutorial to Kai. Here Kai will not mark the further exercises, but will look if an attempt has been made and awards the attendance mark for that week's tutorial.

3.

Let G be a finite group and $x \in G$, $x \neq 1$ (identity element). Define a probability distribution in G:

$$\mu_x = \frac{1}{2}\delta_x + \frac{1}{2}\delta_{-x}$$

Give an example of a finite group G such that the Fourier transform

$$\widehat{\mu_x}(\xi)$$

is unitary for all $\xi \in \widehat{G}$.

Solution 3.

Let $G = \mathbb{Z}_2$. Then we must have $x = 1$ so $x^{-1} = 1$. Hence by definition of the Fourier transform

$$\widehat{\mu_x}(\xi) = \frac{1}{2}\varrho_\xi(x) + \frac{1}{2}\varrho_\xi(x^{-1}) = \frac{1}{2}\varrho_\xi(1) + \frac{1}{2}\varrho_\xi(1) = \varrho_\xi(1),$$

which is unitary for all $\xi \in \widehat{G}$ as the definition of $\widehat{G}$ is the equivalence classes of all unitary (irreducible) representations, so in particular ϱ_ξ is unitary representation.

4.

Recall that the **Uncertainty Principle** in a finite abelian group G said: for all $f : G \to \mathbb{C}$ with $f \neq 0$, we have

$$|\mathrm{spt}(f)||\mathrm{spt}(\widehat{f})| \geq |G|.$$

(we did this in the lecture for $G = \mathbb{Z}_p$, but the proof is same for general abelian G)

Now, prove the following structure theorem relates to this: if G is an abelian group and a function $f : G \to \mathbb{C}$ with $0 \in \operatorname{spt}(f)$ satisfies the equality:

$$|\operatorname{spt}(f)||\operatorname{spt}(\widehat{f})| = |G|,$$

then $\operatorname{spt}(f)$ is a subgroup of G.

Solution 4.

For the proof in full generality see:
https://www.emis.de/journals/AMUC/_vol-73/_no_2/_przebinda/przebinda.pdf
(Matusiak, Özaydin, Przebinda)

Very good supplements for this problem are also:
https://epubs.siam.org/doi/pdf/10.1137/0149053
(Donoho, Stark - This is about the cyclic case)

https://kconrad.math.uconn.edu/blurbs/grouptheory/charthy.pdf
(Conrad)

Let us get some flavor on this. Let us verify this for a very simple case of $\mathbb{Z}_4$. If $|spt(f)| = 1$ or 4, we are done since in these cases $spt(f) = \{0\}$ or $\mathbb{Z}_4$, respectively. Also, $|spt(f)| = 3$ is impossible as then $|spt(\widehat{f})| \notin \mathbb{Z}$. Hence $|spt(f)| = 2$.

- **Case:** $spt(f) = \{0, 1\}$
  Consider:
  $$\widehat{f}(k) = \sum_{a \in \mathbb{Z}_p} f(a) e^{2\pi i k a/4} = f(0) + f(1) e^{2\pi i k/4}.$$
  Clearly $\widehat{f}(1), \widehat{f}(3) \neq 0$. Also both $\widehat{f}(0)$ and $\widehat{f}(2)$ cannot be zero simultaneously. Thus
  $$|spt(f)||spt(\widehat{f})| \neq 4,$$
  this case is not possible.

- **Case:** $spt(f) = \{0, 3\}$
  Similar as the previous case.

- **Case:** $spt(f) = \{0, 2\}$
  Consider:
  $$\widehat{f}(k) = \sum_{a \in \mathbb{Z}_p} f(a) e^{2\pi i k a/4} = f(0) + f(2) e^{\pi i k}.$$
  Here it is possible to have $|spt(\widehat{f})| = 2$.

Hence only suitable and possible case is that $spt(f) = \{0, 2\}$ which is a subgroup. Similar ideas of cancellations (and non-cancellations) should generalize to a more general proof in $\mathbb{Z}_p$.

5.

Construct a probability distribution μ on S_3 with entropy $H(\mu) = \log 2$. Then, using the character table of S_3 and finding the dimensions of the irreducible representations of

S_3, construct some $n_0 \in \mathbb{N}$ such that the entropy $H(\mu^{*n}) > \log 6 + \frac{1}{1000}$ for all $n \geq n_0$.

Solution 5.

$$S_3 = \{e, (12), (13), (23), (123), (132)\}.$$

Let us define μ as follows:

$$\mu = \frac{1}{2}\delta_e + \frac{1}{2}\delta_{(12)}.$$

Let us confirm that it has the wanted entropy:

$$\begin{aligned} H(\mu) &= -\sum_{s \in S_3} \mu(s) log(\mu(s)) \\ &= -\frac{1}{2} log(\frac{1}{2}) - \frac{1}{2} log(\frac{1}{2}) \\ &= \frac{1}{2} log(2) + \frac{1}{2} log(2) \\ &= log(2). \end{aligned}$$

Character table for S_3 is found in the lecture notes from pages 106-107.

We notice that this seems to be wrong! Why? Let $\gamma = \sum_{s \in S_3} a_s \delta_s$ ($a_i \in \mathbb{C}$) be any probability distribution no S_3. Then:

$$\begin{aligned} H(\gamma) &= -\sum_{s \in S_3} \gamma(s) log(\gamma(s)) \\ &= \sum_{s \in S_3} -(\sum_{z \in S_3} a_z \delta_z(s)) log(\sum_{z \in S_3} a_z \delta_z(s)) \\ &\leq [Concavity\ and\ subadditivity] \quad \sum_{s \in S_3} \sum_{z \in S_3} -a_z \delta_z(s) log(a_z \delta_z(s)) \\ &= \sum_{z \in S_3} a_z(-\sum_{s \in S_3} \delta_z(s) log(a_z) - \sum_{s \in S_3} \delta_z(s) log(\delta_z(s))) \\ &= [Latter \ = 0] \sum_{z \in S_3} a_z(-\sum_{s \in S_3} \delta_z(s) log(a_z)) \\ &= \sum_{z \in S_3} -a_z log(a_z) \end{aligned}$$

So we have an optimization problem at hand which we can solve using the method of Lagrange multipliers roughly as follows:

Maximize

$$f(\mathbf{a}) = \sum_{z \in S_3} -a_z log(a_z)$$

with respect to

$$g(\mathbf{a}) = \sum_{z \in S_3} a_z = 1.$$

We see that

$$\nabla g = \mathbf{1}$$

so it being zero matters not. Let

$$L(\mathbf{a}, \lambda) = f(\mathbf{a}) - \lambda g(\mathbf{a}).$$

We want to find suitable values so that
$$\nabla L = \mathbf{0}.$$
We get that for every $z \in S_3$:
$$-log(a_z) - 1 = \lambda \iff a_z = e^{-\lambda-1}.$$
We also get:
$$1 = \sum_{z\in S_3} a_z = \sum_{z\in S_3} e^{-\lambda-1} = 6e^{-\lambda-1} \iff log(\frac{1}{6}) + 1 = -\lambda.$$
Thus $a_z = \frac{1}{6}$. We maximize with $a_z = \frac{1}{6}$. Thus it follows that
$$max(\sum_{z\in S_3} -a_z log(a_z)) = log(6).$$
Clearly we cannot construct the demanded probability distribution.

What was probably the idea of this exercise? First we construct the asked probability distribution with entropy $log(2)$. After this, I think the idea would be to use Pinsker's inequality (**thm.2.34.**) (the proof should work for S_3 as well). Using Pinsker's inequality we would get:
$$\begin{aligned}
& d(\mu^{*n,\lambda}) \geq \frac{1}{2(H(\lambda)+1)}|H(\mu^{*n}) - H(\lambda)| \\
\iff \quad & 2(H(\lambda)+1)d(\mu^{*n},\lambda) \geq H(\lambda) - H(\mu^{*n}) \\
\iff \quad & H(\mu^{*n}) \geq H(\lambda) - 2(H(\lambda)+1)d(\mu^{*n},\lambda). \\
\iff \quad & H(\mu^{*n}) \geq log(6) + 2(log(6)+1)(-d(\mu^{*n},\lambda))
\end{aligned}$$
Entropy for the Lebesgue measure λ can be verified to be $log(6)$. Next we can ask, how big n is needed to get:
$$log(6) - \frac{1}{1000} \leq log(6) + 2(log(6)+1)(-d(\mu^{*n},\lambda))$$
so we would get some estimate on n indicating that
$$H(\mu^{*n}) \geq log(6) - \frac{1}{1000}$$
We proceed. We can use the upper bound lemma to get lowe bound for $(-d(\mu^{*n},\lambda))$:
$$(-d(\mu^{*n},\lambda)) \geq -\frac{1}{2}\sqrt{\sum_{\xi\in\widehat{S_3},\xi\neq 1} dim(V_\xi)||\widehat{\mu}(\xi)^n||^2_{HS,\xi}}.$$
Here, the dimension can be checked from the character table. The Hilbert-Schmidt norm goes like:
$$||\widehat{\mu}(\xi)^n||_{HS,\xi} = \sqrt{Tr_{V_\xi}((\widehat{\mu}(\xi)^n)(\widehat{\mu}(\xi)^n)^*)}.$$
With a good enough choice of μ, the matrix $\widehat{\mu}(\xi)$ should be computable. For this, we should again use the given character table of the lecture notes. After this all is done, perhaps we can compute some n with which we get the wanted property.

In the end, a final note on the fact that with a badly chosen μ, it is possible that μ^{*n} does not converge into λ in total variation distance. Hence, even with a "correct" understanding of the exercise, we might not be able to construct the wanted n_0. Examples of this has been discussed in the earlier exercises.