Introduction to PDEs, Fall 2024

Homework 7 solution

Name:_______________

1. The Sturm–Liouville theory applies to the linear ordinary differential equations of the general form

$$\frac{d}{dx}\left[p(x)\frac{dy}{dx}\right] + q(x)y = -\lambda w(x)y,$$

where p, w are assumed to be of the same sign (typically assumed positive), p, p', q and w are continuous functions of x. If this equation is endowed with the following boundary conditions

$$\begin{aligned} \alpha_1 y(a) + \alpha_2 y'(a) = 0 \quad \alpha_1^2 + \alpha_2^2 > 0, \\ \beta_1 y(b) + \beta_2 y'(b) = 0 \quad \beta_1^2 + \beta_2^2 > 0, \end{aligned}$$

then the statements in our class hold. Let (y_i, λ_i), $i = 1, 2$, be two eigen-pairs of the problem with $\lambda_1 \neq \lambda_2$. Prove that y_i are orthogonal in (weighted) L^2 as

$$\langle y_1, y_2 \rangle := \int_a^b y_1(x) y_2(x) w(x) dx = 0.$$

Remark: Indeed, one can show that

$$\langle y_n, y_m \rangle = \int_a^b y_n(x) y_m(x) w(x) dx = \delta_{mn},$$

where δ_{mn} is the Kronecker delta function.

Solution 1. *The proof is the same as those for the case when $p, w \equiv 1$ and $q \equiv 0$. One tests the y_1 equation by y_2, the y_2 equation by y_1, apply the integration by parts, and then using the boundary conditions to show that the endpoint values are cancelled. I skip the rest as they are routine calculations. What I want to emphasize is that this conclusion is practically important and useful.*

2. Some of you may be wondering why we care about and look for a solution (only) in L^2. This is a very good question, and here are some of the main reasons. First of all, it is obvious that a solution in L^2 is not enough, and ultimately we are interested in finding smooth solutions. However, you will find that you lack the tools to do this at this stage. Therefore, one can first find some solutions in L^2, and then prove (if we can and if they are) that they are actually in L^4, L^6,...L^∞ or other spaces (Sobolev spaces), and then finally smooth (or not) by techniques called *a prior estimate* or *regularity estimate*. Of course, other advanced mathematics must come into play, and they are beyond the scope of this course. For example, one of the seven Millennium Prize Problems in Mathematics proposed by the Clay Mathematics Institute in May 2000 is to prove the existence or nonexistence of a smooth and globally well-defined solution to the Navier-Stokes equations in 3D, but it remains open to this day. Second, some equations (e.g., wave equations) do not have smooth solutions and can develop a shock or singularity in finite time. Many reaction-diffusion equations/systems also admit solutions with their L^∞ going to infinity in finite time, and this is called blow-up in finite time. Therefore, finding a smooth solution for all time $t \in \mathbb{R}^+$, is impossible for many problems. I also want to mention that saying that L^2 is not enough does not mean that it is not interesting, people nowadays are still very interested in finding solutions in L^p spaces, or the so-called weak solutions, but the method of separation of variables does not work very well, not only because of the PDE, but also because of the properties of the space itself, especially if the function (solution) has a jump discontinuity. The eigen–expansion or Fourier expansion of $f(x)$ is the limit of the following sum as $N \to \infty$

$$f^N(x) := \sum_{n=0}^{\infty} C_n X_n(x),$$

where $\{X_n(x)\}$ is the orthogonal basis of $L^2(0,L)$ and

$$C_n = \frac{\int_0^L f(x)X_n(x)dx}{\int_0^L X_n^2(x)dx}.$$

then the general theory states that $f^N \to f$ pointwise in $(0,L)$ if $f(x)$ is continuous, uniformly if $f(x)$ is differentiable, while in L^2 if f is merely square integrable. In the last case, in particular, when $f(x)$ has jump discontinuity, one may see that the approximation can go wild, and this is usually referred to as the so-called *Gibbs phenomenon.* To see this yourself, let us consider the following example: let $f(x)$ be a function with a jump at $x = 0$ defined over $(-\pi, \pi)$ as: $f(x) = 1$ for $x \in [0, \pi)$ and $f(x) = -1$ for $x \in (-\pi, 0)$. First of all, write $f(x)$ into its series as

$$f(x) = \sum_{n=1}^{\infty} C_n \sin nx. \tag{0.1}$$

Find C_n. Now let us approximate it by the sum of the first N terms as before

$$f_N(x) := \sum_{n=1}^{N} C_n \sin nx \tag{0.2}$$

for some large N. Plot $f_N(x)$ over $(-\pi, \pi)$ for $N = 2, 4, 8, 16$ on the same graph. Try $N = 16, 32$ and 64 again. What are your observations? You can try with even larger N.

Solution 2. *I skip the calculations and $C_n = \frac{2}{n}(1-(-1)^n)$. The Gibbs phenomenon is evident from the plot.*

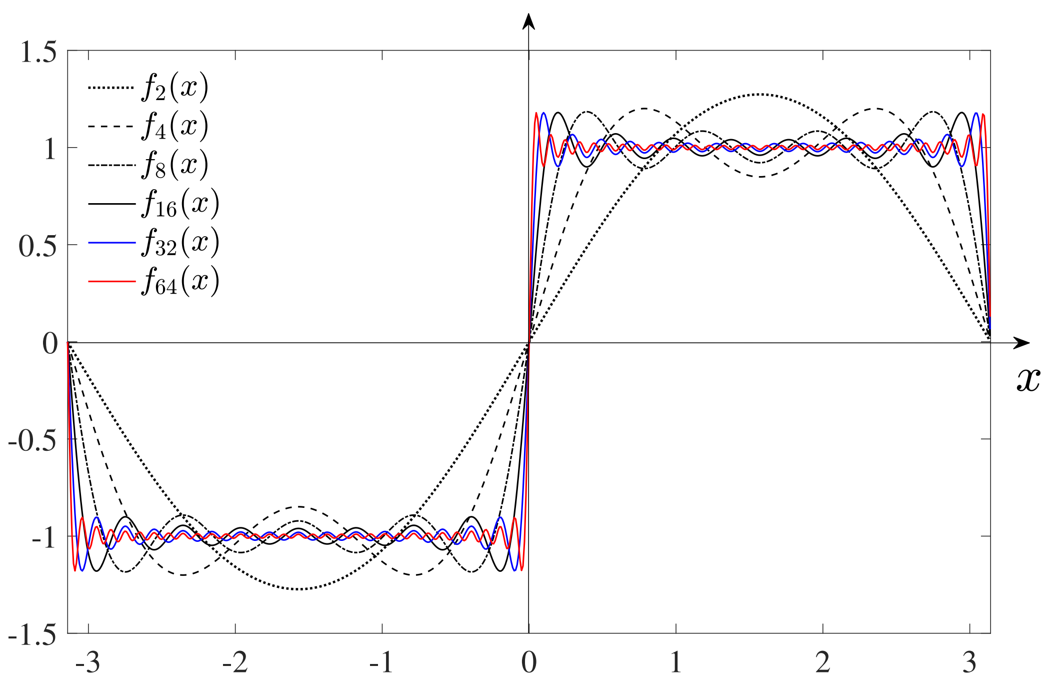


Figure 1: Finite sum f^N and oscillations for N large. We observe the non-convergence (in a pointwise sense) of f^N for N large, though the error converges to zero in L^2. Again this is because its limit is merely L^2, but not continuous, hence pointwise convergence is not expected.

3. We recall from baby calculus that a number sequence is called a Cauchy sequence if $|a_{n+1} - a_n|$ is arbitrarily small if n is arbitrarily large; for example $a_n = \frac{1}{n}$ is Cauchy because $|a_{n+1} - a_n| = \frac{1}{n(n+1)} \to 0$ as $n \to \infty$. And a region is called **complete** if every Cauchy converges in it. For example, $(0, 1)$ is not closed since $a_n = \frac{1}{n}$ is Cauchy but $a_n \to 0 \notin (0, 1)$, while $[0, 1]$ is closed. There are more or less rigorous definitions of a closed region. We have a cousin of closedness when it comes to function space, i.e., completeness. For a function space $\mathcal{X}$ endowed with a norm (called a normed space or a metric space) denoted by $\|\cdot\|_{\mathcal{X}}$, we say that a function

sequence is Cauchy if $\|f_{n+1}(x)-f_n(x)\|_{\mathcal{X}} \to 0$ as $n \to \infty$. Here the only difference/generalization is that the absolute value distance is replaced by the so-called "norm"; then we say that space $\mathcal{X}$ is **complete** if every Cauchy sequence converges (in the space or that norm). For instance, one can prove that $L^p(\Omega)$, $p \in (1,\infty)$ is complete (well, you should have learned in your Analysis course or learned it by yourself otherwise); moreover, in Euclidean space complete is equivalent as closed and bounded.

Going back to the completeness of L^p, if f_n is Cauchy in L^p its limit must be in L^p. Then one usually needs to verify the Cauchyness of the sequences.

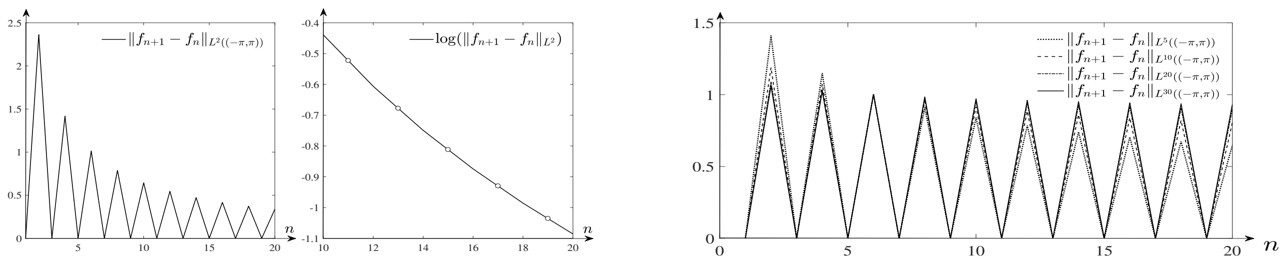
i) prove that the sequence f_N given by (0.2) is Cauchy in L^2. Then according to i), its limit (0.1) belongs to L^2 and this gives you proof of the convergence fashion we had in class;

ii) plot $\|f_{n+1}-f_n\|_{L^2((-\pi,\pi))}$ for $n=1,2,\ldots$ and you should observe the convergence; indeed, to better illustrate the convergence it makes sense to plot the log error $\mathcal{E}_n := \log(\|f_{n+1}-f_n\|_{L^2})$ for n large, for instance starting from 10. Remark: MATLAB can evaluate the integrals hence you do not have to do it by brutal force;

iii) try different L^p norms with $p = 5, 10, 20, 30, \ldots$ and do the same as in ii); plot them in the same graph; what are your observations?

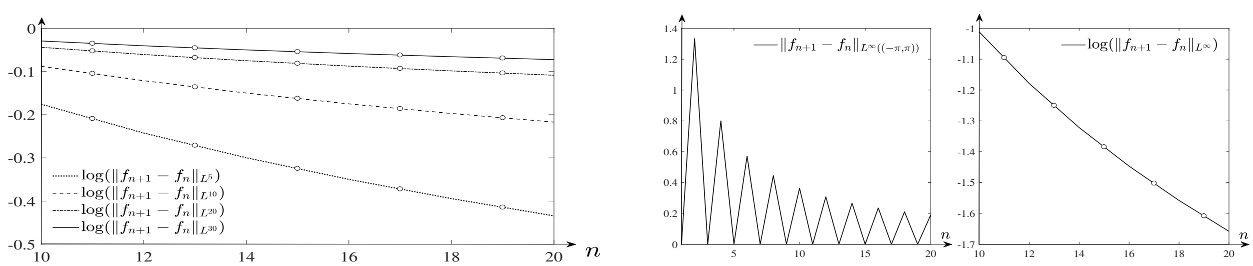
iv) now find the max-norm and do the same as in ii).

Solution 3. *The verification of Cauchyness is rather straightforward and I skip the calculations here.*



Absolute errors in L^2 | Absolute errors in L^p

Figure 2



Log errors in L^p for different p | Errors in L^∞ and log errors

Figure 3: For each p one observes the oscillating dissipation of errors. It seems that an exponential decay is not expected until very large N.

4. This problem is to warm up and prepare you for the fashions of convergence of sequences of functions. The formal way (using ϵ–δ language) to define **pointwise convergence** of $f_n(x)$ over set E is, $\forall x \in E$ and $\forall \epsilon > 0$, there exists $N_0 \in \mathbb{N}^+$ such that for all $n \geq N_0$ we have that $|f_n(x)-f(x)| \leq \epsilon$. **Uniform convergence** of $f_n \to f$ is that, $\forall \epsilon > 0$, there exists $N_1 \in \mathbb{N}^+$ such that for all $n \geq N_1$ we have that $|f_n(x)-f(x)| \leq \epsilon$ $\forall x$ in E. The difference is the order of $\forall x$ and N, therefore N_0 depends on both ϵ and x, while N_1 does not depend on x hence by saying uniform convergence we meant that it is uniform with respect to x, or the convergence speed does not depend on x.

   To elaborate, we recall that a function $f(x)$ is **continuous** at x_0 if $\forall \epsilon > 0$, we can find $\delta > 0$ such that $|x-x_0| < \delta$ implies $|f(x)-f(x_0)| < \epsilon$. Here δ depends on ϵ and also x_0. If δ only depends on ϵ and is independent of x_0, i.e., $|x-y| < \delta$ implies $|f(x)-f(y)| < \epsilon$ for any $x, y \in E$, then f is **uniformly**

continuous. There is another type of continuity called **equicontinuity** for a sequence of functions $\{f_n(x)\}$, which states that $|x - x_0| < \delta$ implies $|f_n(x) - f_n(x_0)| < \epsilon$ for all n, while δ may depend on x_0 but does not depend on n. Similarly, if δ only depends on ϵ, we say that $\{f_n\}$ is **uniform equicontinuous**. An important property of equicontinuous function sequence according to **Arzela-Ascoli** Theorem is that any uniformly bounded and equicontinuous sequence has a subsequence that converges uniformly. This theorem is one of two cornerstones in modern mathematical analysis (the other being semi-continuity), and you are not required to understand equicontinuity for this course.

We already know that uniform convergence implies pointwise convergence but not the other way. However, if the limit is also continuous, then pointwise convergence implies uniform convergence (You might be asked to prove this in an advanced analysis course).

(a). Prove that if a sequence of continuous functions f_n converges to f uniformly in (a, b), then f is continuous. (Hint: use the $\epsilon - \delta$ language; for each x_0 in E, choose $f_n(x_0)$ that converges to $f(x_0)$; do the same for $f_n(x)$ with x being in the neighbourhood of x_0.)

(b). Another important application of uniform convergence is the switch of the order of integral and limit. Consider the following so-called tent–function

$$f_n(x) = \begin{cases} n^2 x, & x \in [0, \frac{1}{n}], \\ 2n - n^2 x, & x \in (\frac{1}{n}, \frac{2}{n}], \\ 0, & x \in (\frac{2}{n}, 1], \end{cases} \tag{0.3}$$

Find the pointwise limit $f(x)$ of $f_n(x)$; evaluate the limits

$$\lim_{n\to\infty} \int_0^1 f_n(x)dx \text{ and } \int_0^1 \lim_{n\to\infty} f_n(x)dx.$$

Are they equal?

(c). What is Lebesgue's Dominated Convergence Theorem? Why does it fail in the example above?

(d). Let $f_n(x)$ be a sequence that converges to $f(x)$ uniformly in interval (a, b). Prove that

$$\lim_{n\to\infty} \int_a^b f_n(x)dx = \int_a^b \lim_{n\to\infty} f_n(x)dx = \int_a^b f(x)dx.$$

Solution 4. *(a). To show that f is continuous at any point $x_0 \in (a, b)$, we want to show that $\forall \epsilon > 0$, there exists $\delta > 0$ (which may depend on x_0) such that*

$$f(x) - f(x_0) \le \epsilon, \forall x \in B_\delta(x_0),$$

where $B_\delta(x_0)$ in general means a small NBHD of x_0 with radius δ and here is just $B_\delta(x_0) = \{x; |x - x_0| < \delta\}$. To prove this, in light of uniform convergence of f_n to f, we can always find N_0, which depends on ϵ but not x such that for all $n \ge N_0$ we have

$$|f_n(x) - f(x)| \le \frac{\epsilon}{3} \forall x \in B_\delta(x_0);$$

in particular, this also holds when $x = x_0$, i.e., for all $n \ge N_0$

$$|f_n(x_0) - f(x_0)| \le \frac{\epsilon}{3}.$$

On the other hand, since $f_n(x)$ is continuous for each n, we choose $n = N_0 + 1$ and have that

$$f_{N_0+1}(x) - f_{N_0+1}(x_0) \le \frac{\epsilon}{3} \forall x \in B_\delta(x_0),$$

and therefore, by choosing $n = N_0 + 1$ above we can finally have that

$$|f(x) - f(x_0)| \le |f(x) - f_{N_0+1}(x)| + |f_{N_0+1}(x) - f_{N_0+1}(x_0)| + |f_{N_0+1}(x_0) - f(x_0)| \le \epsilon,$$

which is exactly what we have desired.

***Remark:** This result is known as the Uniform Limit Theorem. The statement remains valid if continuity is replaced by uniform continuity. Specifically, if f_n is a sequence of uniformly continuous functions that converge uniformly to f, then f must also be uniformly continuous. You are encouraged to try proving this yourself.*

(b). The pointwise limit of f_n is $f(x) \equiv 0$, $x \in [0,1]$. To see this, you should fix each $x \in [0,1]$ and then send $n \to \infty$, therefore we have that

$$\int_0^1 f(x)dx = 0;$$

however, on the other hand

$$\int_0^1 f_n(x)dx = 1$$

hence

$$\lim_{n\to} \int_0^1 f_n(x)dx = 1,$$

therefore the limit of the integral and integral of the limit do not equal, i.e., the order of the limit and integration can not be switched.

*(c). The **Lebesgue Dominated Convergence Theorem (LDCT)** states that if $f_n \to f$ pointwise (or almost everywhere, meaning the set of points where it does not converge has measure zero) and $|f_n| \leq M$ for some constant $M > 0$ independent of n, then:*

$$\lim_{n\to\infty} \int_\Omega f_n(x)\, dx = \int_\Omega f(x)\, dx.$$

The theorem fails in the example above because the tent function is not uniformly bounded. Without the uniform boundedness condition, the integrability required for LDCT may break down, preventing the limit from being interchanged with the integral.

(d). However, if f_n converges to f uniformly, the order of integration and the limit can be interchanged. To see this, note that uniform convergence implies that for any $\epsilon > 0$, there exists N such that

$$|f_n(x) - f(x)| < \frac{\epsilon}{b-a}, \quad \forall x \in (a,b),$$

for all $n \geq N$. Therefore, we have

$$\left| \int_a^b f_n(x)\, dx - \int_a^b f(x)\, dx \right| \leq \int_a^b |f_n(x) - f(x)|\, dx \leq \frac{\epsilon}{b-a} \cdot (b-a) = \epsilon.$$

This demonstrates the expected result.

5. It is known that strong convergence implies weak convergence, while not the converse. One counter-example we mentioned in class is $f_n(x) := \sin nx$ over $(0, 2\pi)$ (or $(0,\pi)$).

   (i) Prove that $\sin nx \nrightarrow 0$ in $L^2((0,2\pi))$.

   (ii) Prove that $\sin nx \rightharpoonup 0$ weakly by showing

   $$\int_0^{2\pi} g(x) \sin nx dx \to 0 = \Big(\int_0^{2\pi} g(x) 0 dx \Big) \quad \forall g \in L^2((0,2\pi)).$$

   If suffices even if $g \in L^1$. Hint: Riemann–Lebesgue lemma.

   (iii). Find another counter-example in a textbook or online.

   Solution 5. *skipped.*

6. (**only for motivated students**) We recall that $f_n(x) \rightharpoonup f(x)$ weakly in L^p (resp. convergence in distribution) if for any $\phi \in L^q$ (resp. continuous and bounded), which is its conjugate space with $\frac{1}{p} + \frac{1}{q} = 1$, we have that

   $$\int_\Omega f_n \phi dx \to \int_\Omega f \phi dx.$$

Here we see that for any g in L^q

$$< \cdot, g >= \int_\Omega \cdot g$$

defines a bounded linear functional for L^p. Then we also call L^q *the dual space* of L^p since any element in L^q defines a functional for L^q.

(i) Another type of convergence that you may see sometimes is $\|f_n\|_p \to \|f\|_p$, which merely states the convergence of a sequence of real numbers. Prove that if $f_n \to f$ in L^p, then $\|f_n\|_p \to \|f\|_p$ (Use Minkowski triangle inequality); however the opposite statement is not necessarily true. Give a counter-example and show it;

(ii) Prove that, if $f_n \rightharpoonup f$ weakly and $\|f_n\|_p \to \|f\|_p$, then $f_n \to f$ strongly.

Solution 6. *Minkowski's triangle inequality states that for any $f, g \in L^p$ with $p \in (1, \infty)$, we have:*

$$\|f + g\|_{L^p} \leq \|f\|_{L^p} + \|g\|_{L^p}.$$

From this, it is straightforward to deduce:

$$\big|\|f_n\|_{L^p} - \|f\|_{L^p}\big| \leq \|f_n - f\|_{L^p}.$$

If $\|f_n - f\|_{L^p} \to 0$, then $\|f_n\|_{L^p} \to \|f\|_{L^p}$, which proves the desired claim. You can work through the detailed derivation on your own.

I will skip presenting a specific counterexample, but here's a hint for constructing one: Think of f_n and f as points in a plane, where their norms represent distances from the origin. The condition $\|f_n - f\|_{L^p} \to 0$ implies that f_n approaches f (the distance between them goes to zero). However, $\|f_n\|_{L^p} \to \|f\|_{L^p}$ merely implies that the distance of f_n from the origin converges to that of f. It's easy to see that the former implies the latter, but not vice versa. With this reasoning, you should be able to construct a suitable counterexample.

Finally, we shall prove that the two conditions can imply a strong convergence. To prove this, let us divide our discussions into the following cases:

case 1: $p = 2$. Then the conclusion is straightforward following Cauchy–Schwarz inequality. I assume that you have no problem proving this case;

case 2: $p > 2$. We first see that for any $z \in \mathbb{R}$

$$|z + 1|^p \geq c|z|^p + pz + 1,$$

where c is a positive constant independent of z. (To prove this fact, we just need to show that $\frac{|z+1|^p - pz - 1}{|z|^p}$ has a positive lower bounded c over $\mathbb{R}$). Now we can let $z = \frac{f_n - f}{f}$ in this inequality, multiply it by $|f|^p$ and then integrate the new one over Ω to obtain

$$\int_\Omega |f_n|^p dx \geq \int_\Omega |f|^p dx + p \int_\Omega |f|^{p-2} f(f_n - f) dx + c \int_\Omega |f_n - f|^p dx.$$

Since $f_n \rightharpoonup f$ weakly, we see that the second integral on the right-hand side of the equality converges to zero (think of $|f|^{p-2}f$ as a test function). On the other hand, we have that $\int_\Omega |f_n|^p dx \to \int_\Omega |f|^p dx$ thanks to the strong convergence, therefore we must have

$$\int_\Omega |f_n - f|^p dx \to 0,$$

which implies a strong convergence.

case 3: $p \in (1, 2)$. The proof of this part is a little bit tricky. Similar to above, we can show (by straightforward calculations) that $\forall z \in \mathbb{R}$

$$\begin{aligned} |z + 1| \geq & c|z|^p + pz + 1, \ \text{if } |z| \geq 1, \\ |z + 1| \geq & c|z|^2 + pz + 1, \ \text{if } |z| \geq 1. \end{aligned}$$

In order to apply these inequalities, we shall choose $z = \frac{f_n - f}{f}$. Denote

$$\Omega_n := \{x \in \Omega; |z| \geq 1\},$$

then we can have by the same calculations as above that

$$\begin{aligned}\int_\Omega |f_n|^p dx &= \int_{\Omega\backslash\Omega_n} |f_n|^p dx + \int_{\Omega_n} |f_n|^p dx\\ &= \int_{\Omega\backslash\Omega_n} |z+1|^p |f|^p dx + \int_{\Omega_n} |z+1|^p|f|^p dx\\ &\geq \int_{\Omega\backslash\Omega_n} (c|z|^2+pz+1)|f|^p dx + \int_{\Omega_n} (c|z|^p+pz+1)|f|^p dx,\end{aligned}$$

which implies, in light of the formula of z, that

$$\int_{\Omega\backslash\Omega_n} (f_n-f)^2|f|^{p-2}dx + \int_{\Omega_n} |f_n-f|^p dx \to 0.$$

Both integrals are nonnegative, hence both should converge to zero

$$\int_{\Omega\backslash\Omega_n} (f_n-f)^2|f|^{p-2}dx \to 0, \int_{\Omega_n} |f_n-f|^p dx \to 0.$$

In particular, we only need to show that

$$\int_{\Omega\backslash\Omega_n} |f_n-f|^p dx \to 0.$$

For this purpose, we shall apply Holder's inequality or Schwarz's inequality as follows. Note that $|f_n - f| < |f|$ in $\Omega\backslash\Omega_n$, then we have that

$$\begin{aligned}\int_{\Omega\backslash\Omega_n} |f_n-f|^p dx &\leq \int_{\Omega\backslash\Omega_n} |f|^{p-1}|f_n-f|dx\\ &\leq \Big(\int_{\Omega\backslash\Omega_n} |f|^p\Big)^{\frac{1}{2}}\Big(\int_{\Omega\backslash\Omega_n} |f|^{p-2}|f_n-f|^2dx\Big)^{\frac{1}{2}}\\ &\leq \Big(\int_{\Omega} |f|^p\Big)^{\frac{1}{2}}\Big(\int_{\Omega\backslash\Omega_n} |f|^{p-2}|f_n-f|^2dx\Big)^{\frac{1}{2}} \to 0,\end{aligned}$$

which is the desired claim and the proof completes.