Second order elliptic equations with a small parameter.

Mark Freidlin $^1$, Wenqing Hu$^2$.

$^1$Department of Mathematics, University of Maryland, College Park.
$^2$Department of Mathematics, University of Maryland, College Park.
We are interested in boundary problems for the operator \( L_\varepsilon = L_0 + \varepsilon L_1 \) and initial-boundary problems for \( \frac{\partial u^\varepsilon}{\partial t} = L_\varepsilon u^\varepsilon \), \( t > 0, \ x \in \partial G \).

Operators

\[
L_k = \mathbf{b}^{(k)}(x) \cdot \nabla + \frac{1}{2} \sum_{i,j=1}^{d} a^{(k)}_{ij}(x) \frac{\partial^2}{\partial x_i \partial x_j}, \quad k = 0, 1.
\]

Vectors \( \mathbf{b}^{(k)}(x) = (b^{(k)}_1(x), \ldots, b^{(k)}_d(x)), \quad k = 0, 1. \)

Coefficients \( a^{(k)}_{ij} \) and \( b^{(k)}_j \) are \( \mathcal{C}^2 \).

For fixed \( \varepsilon > 0 \) the operator \( L_\varepsilon \) is elliptic.
We take the Dirichlet problem as an example:

\[ L_\varepsilon u_\varepsilon(x) = (L_0 + \varepsilon L_1)u_\varepsilon(x) = 0, \quad u_\varepsilon(x)|_{\partial G} = \psi(x). \]

If \( L_0 \) is elliptic, then \( u_\varepsilon \to u^0 \) as \( \varepsilon \downarrow 0 \) where \( L_0 u^0(x) = 0 \), \( u^0(x)|_{\partial G} = \psi(x) \).

What about degenerate \( L_0 \)?
Levinson case.

- Levinson case (1950):

  \[ L_0 = b^{(0)}(x) \cdot \nabla \]

  and

  \[ L_1 = b^{(1)}(x) \cdot \nabla + \frac{1}{2} \sum_{i,j=1}^{d} a_{ij}^{(1)}(x) \frac{\partial^2}{\partial x_i \partial x_j} \]

- Levinson condition: trajectories of the dynamical system \( \dot{X}_t = b^{(0)}(X_t) \) leave the domain \( G \) in finite time and cross the boundary \( \partial G \) in a regular way.

- Theorem. (Levinson, 1950) We have \( \lim_{\varepsilon \downarrow 0} u^\varepsilon(x) = u^0(x) \) where \( u^0(x) \) is the solution of \( L_0 u^0(x) = 0, \ x \in G \) and \( u^0|_{\partial_1 G} = \psi(x) \). Here \( \partial_1 G \) is the part of the boundary \( \partial G \) where \( X_t \) hits and leaves \( \partial G \).
FIG.: Levinson case.
Levinson case from probabilistic point of view.

- Levinson’s result can be easily explained from a probabilistic point of view.
- Let as before $\dot{X}_t = b^{(0)}(X_t)$.
- Consider a diffusion process with a small diffusion:

$$dX_t^\varepsilon = \sqrt{\varepsilon}\sigma^{(1)}(X_t^\varepsilon)\,dW_t + b^{(0)}(X_t^\varepsilon)\,dt,$$

$$X_0^\varepsilon = x \in \mathbb{R}^d.$$

- $\sigma^{(1)}(x)(\sigma^{(1)}(x))^* = a^{(1)}(x)$.
- As $\varepsilon \downarrow 0$ the process $X_t^\varepsilon$ converges to $X_t$ in a certain sense.
- Thus $u^\varepsilon(x) = \mathbf{E}_x\psi(X_t^\varepsilon) \to \mathbf{E}_x\psi(X_t) = u^0(x)$ as $\varepsilon \downarrow 0$.
- $\tau$ is the first hitting time of $X_t^\varepsilon$ to $\partial G$. 
Fig.: Levinson case.
Summary.

- **Summary**: Convergence of the solution of corresponding PDE $\iff$ (Weak) convergence of corresponding diffusion process.
Degenerate problems : Neumann case.

- Neumann problem
  \[
  \left( \frac{1}{\varepsilon}L_0 + L_1 \right) u^\varepsilon(x) = f(x), \quad \frac{\partial u^\varepsilon(x)}{\partial \gamma^\varepsilon(x)}\bigg|_{\partial G} = 0.
  \]

- We work with self-adjoint situation
  \[
  L_k u(x) = \frac{1}{2} \nabla \cdot (a^{(k)}(x) \nabla u(x)).
  \]

- Solvability and uniqueness condition : for Hölder continuous
  \[
  \int f(x) \, dx = 0 \quad \text{and for } x_O \in G \cup \partial G \text{ we have } u(x_O) = 0.
  \]

- \(a^{(0)}(x)\) is degenerate and non-negative definite; \(a^{(1)}(x)\) is positive definite. Suppose \(a^{(0)}(x) = \sigma^{(0)}(x)(\sigma^{(0)}(x))^*\) and \(a^{(1)}(x) = \sigma^{(1)}(x)(\sigma^{(1)}(x))^*\). The coefficients of \(a^{(1)}(x)\) are in \(C^{(2)}\) and the coefficients of \(a^{(0)}(x)\) are in at least \(C^{(1)}\).

- We specify the degeneration by looking at a first integral
  \(H(x) : a^{(0)}(x) \nabla H(x) = 0\).

- We single out only one first integral by making a restriction
  \[e \cdot (a^{(0)}(x)e) \geq a(x)|e|^2_{\mathbb{R}^d} \quad \text{for each } e \text{ such that } e \cdot \nabla H(x) = 0.\]
First we assume that $a^{(0)}(x)$ has constant rank $d-1$ and its coefficients are in $C^{(2)}$.

Corresponding process

$$
dX_t^\varepsilon = \frac{1}{\varepsilon} b^{(0)}(X_t^\varepsilon) dt + \frac{1}{\sqrt{\varepsilon}} \sigma^{(0)}(X_t^\varepsilon) dW^0_t \quad \text{(fast motion)}
$$

$$
+ b^{(1)}(X_t^\varepsilon) dt + \sigma^{(1)}(X_t^\varepsilon) dW^1_t.
$$

with reflection w.r.t. inward co-normal $\gamma^\varepsilon(x)$ at $\partial G$.

$b^{(0)}$ and $b^{(1)}$ are calculated from $a^{(0)}$ and $a^{(1)}$.

Fast motion is moving on level surface $\{H(x) = \text{const}\}$ and has Lebesgue measure as its invariant measure.

Averaging principle (Khasminski, Freidlin-Wentzell, ...) : The limit of slow motion $H(X_t^\varepsilon)$ can be calculated by averaging with respect to the fast motion : $H(X_t^\varepsilon) \to Y_t$ weakly as $\varepsilon \downarrow 0$. 
Fig.: Averaging Principle.
Averaging principle.

- $H(X^\varepsilon_t) \to Y_t$ weakly in $C_{[0,T]}(\Gamma)$.
- $Y_t$ is a 1-dimensional process.
- In the simplest case as in our example $Y_t$ has a generator
  \[ \mathcal{L}f(H) = \frac{1}{2} M^{-1}(H) \frac{d}{dH} \left( M(H) a^{(1)}(H) \frac{df}{dH} \right). \]
- Here $a^{(1)}(h) = M^{-1}(h) \int_{C(h)} \frac{(a^{(1)}(x) \nabla H(x), \nabla H(x))}{|\nabla H(x)|_{\mathbb{R}^d}} d\sigma$ and $M(h) = \int_{C(h)} \frac{d\sigma}{|\nabla H(x)|_{\mathbb{R}^d}}$. 

\[ H(X^\varepsilon_t) \to Y_t \text{ weakly in } C_{[0,T]}(\Gamma). \]
\[ Y_t \text{ is a 1-dimensional process.} \]
\[ \text{In the simplest case as in our example } Y_t \text{ has a generator} \]
\[ \mathcal{L}f(H) = \frac{1}{2} M^{-1}(H) \frac{d}{dH} \left( M(H) a^{(1)}(H) \frac{df}{dH} \right). \]
\[ \text{Here } a^{(1)}(h) = M^{-1}(h) \int_{C(h)} \frac{(a^{(1)}(x) \nabla H(x), \nabla H(x))}{|\nabla H(x)|_{\mathbb{R}^d}} d\sigma \text{ and} \]
\[ M(h) = \int_{C(h)} \frac{d\sigma}{|\nabla H(x)|_{\mathbb{R}^d}}. \]
The implication of averaging principle on differential equations.

▶ Neumann problem:
\[
\left( \frac{1}{\varepsilon}L_0 + L_1 \right) u^\varepsilon(x) = f(x) \quad \text{for } x \in G, \quad \frac{\partial u^\varepsilon(x)}{\partial \gamma^\varepsilon(x)} \bigg|_{x \in \partial G} = 0.
\]

▶
\[
u^\varepsilon(x) = - \int_0^\infty E_x f(X_t^\varepsilon) dt + \int_0^\infty E_{x_0} f(X_t^\varepsilon) dt.
\]
Fig.: Identification mapping $\mathcal{Y}$. 
The implication of averaging principle on differential equations.

\[ u^\varepsilon(x) = -\int_0^\infty \mathbf{E}_x f(X^\varepsilon_t)dt + \int_0^\infty \mathbf{E}_{x_0} f(X^\varepsilon_t)dt \]
\[ \to -\int_0^\infty \mathbf{E}_{2y(x)} \bar{f}(Y_t)dt + \int_0^\infty \mathbf{E}_{2y(x_0)} \bar{f}(Y_t)dt = v(2y(x)) . \]

\[ \bar{f}(h) = \frac{1}{M(h)} \int_{C(h)} f(x) \frac{d\sigma}{|\nabla H(x)|_{\mathbb{R}^d}} . \]

\[ v(h) \text{ is the solution of ODE on } \Gamma : \mathcal{L}v(h) = -\bar{f}(y) \text{ and } v(2y(x_0)) = 0. \]
A few remarks.

- The case when $H(x)$ has saddle point in $G$ : gluing condition at the interior vertices of $\Gamma$. (Freidlin-Wentzell, PTRF, 2012)
- The case of attractor : In the sense of random perturbations of dynamical systems – large deviation principle. (Freidlin-Wentzell, 1969 ; Kifer, 1974)
Our problem.

- We are interested in the case that $a^{(0)}(x)$ has rank $d$ in $\mathcal{E} \subset G$ and rank $d - 1$ in $[G] \setminus \mathcal{E} = \bigcup_{k=1}^{r} [U_k]$.
- Global first integral $H(x): H(x) = 0$ on $[\mathcal{E}]$ and $H(x) = H_k(x)$ on $[U_k]$. 
FIG.: Our problem.
Assumption on the degeneracy.

- On \( \bigcup_{k=1}^{r} [U_k] \) : \( a^{(0)}(x) \) has rank \( d - 1 \). Existence of first integrals \( H_k, \ k = 1, \ldots, r \). Non-dengeneracy on \( C_k(h) = \{ x \in U_k : H_k(x) = h \} \), etc. similar assumptions as in the averaging principle.

- Denote \( \gamma_k = \partial U_k \) and \( \gamma = \bigcup_{k=1}^{r} \gamma_k \).

- \( \text{const}_1 \cdot \text{dist}^2(x, \gamma) \leq e_d(x) \cdot (a^{(0)}(x)e_d(x)) \leq \text{const}_2 \cdot \text{dist}^2(x, \gamma) \). ("quadratic degeneracy")

- Coefficients of \( a^{(0)}(x) \) are in \( \mathbf{C}^{(1)} \) on \( \gamma \). We assume the decomposition \( \sigma^{(0)}(x)(\sigma^{(0)}(x))^* = a^{(0)}(x) \).
**Figure:** Identification mapping \( \mathcal{Y} \).
Weak convergence of the process $Y^\varepsilon_t = \mathcal{Y}(X^\varepsilon_t)$.

- Let $Y^\varepsilon_t = \mathcal{Y}(X^\varepsilon_t)$.
- We can introduce the graph $\Gamma$ with coordinate $(k, H)$.
- $Y^\varepsilon_t$ lives on $\Gamma$ and is, in general, not Markov for fixed $\varepsilon > 0$.
- $Y^\varepsilon_t \to Y_t$ weakly as $\varepsilon \downarrow 0$ in $C[0,T](\Gamma)$.
- Inside each $I_k$ the process $Y_t$ has a generator

$$
\mathcal{L}_k f(k, H_k) = \frac{1}{2} M_k^{-1}(H_k) \frac{d}{dH_k} \left( M_k(H_k) a^{(1)}(H_k) \frac{df}{dH_k} \right).
$$

- Here

$$
\overline{a^{(1)}}(h) = M_k^{-1}(h) \int_{C_k(h)} \frac{(a^{(1)}(x) \nabla H_k(x), \nabla H_k(x))}{|\nabla H_k(x)|_{\mathbb{R}^d}} d\sigma,
$$

and normalizing factor

$$
M_k(h) = \int_{C_k(h)} \frac{d\sigma}{|\nabla H_k(x)|_{\mathbb{R}^d}}.
$$
Weak convergence of the process $Y^\varepsilon_t = \mathcal{Y}(X^\varepsilon_t)$.

- What is more important: $Y_t$ is a Markov process on $\Gamma$ with generator $A$ and domain of definition $D(A)$.
- Inside each $I_k$ $A$ agrees with $\mathcal{L}_k$. (standard averaging principle)
- The domain of definition $D(A)$ consists of those functions $f$ that are twice continuously differentiable inside each $I_k$ having the limit $\lim_{H_k \to 0} \frac{\partial f}{\partial H_k}(k, H_k)$. These functions satisfy the gluing condition at the vertex $O$:

$$0 = \text{Volume}(\mathcal{E}) \cdot Af(O) + \frac{1}{2} \sum_{k=1}^r p_k \cdot \lim_{H_k \to 0} \frac{\partial f}{\partial H_k}(k, H_k).$$

(gluing condition of "delay" type)

- Here $\text{Volume}(\mathcal{E})$ is the $d$-dimensional volume of the domain $\mathcal{E}$ and

$$p_k = \int_{\gamma_k} \frac{(a^{(1)}(x) \nabla H_k(x), \nabla H_k(x))}{|\nabla H_k(x)|_{\mathbb{R}^d}} d\sigma.$$

Fig.: gluing condition of "delay" type.
The answer to our problem.

- **Theorem.** (Freidlin-H, 2012, preprint) *Consider the Neumann problem*

\[
\frac{1}{\varepsilon} L_\varepsilon u_\varepsilon(x) = \left( \frac{1}{\varepsilon} L_0 + L_1 \right) u_\varepsilon(x) = f(x) \text{ for } x \in G ,
\]

\[
\frac{\partial u_\varepsilon(x)}{\partial \gamma_\varepsilon(x)} \bigg|_{x \in \partial G} = 0
\]

*with a Hölder continuous function* \( f(x) \)* satisfying

\[
\int_G f(x) dx = 0. \text{ Let } u_\varepsilon(x_0) = 0 \text{ for some } x_0 \in G. \text{ Then we have}
\]

\[
\lim_{\varepsilon \downarrow 0} u_\varepsilon(x) = v(\gamma(x))
\]

*where* \( v(y) \)* is a continuous function on* \( \Gamma \)* such that

\[
\mathcal{L}_k v(y) = -\bar{f}(y) \text{ for } y \in (l_k) , \ k = 1, \ldots, r .
\]
The answer to our problem.

▶ **Theorem.** (continued) (Freidlin-H, 2012, preprint)  

Here

\[
\bar{f}(y) = \frac{1}{\text{Volume}(\mathcal{E})} \int_{\mathcal{E}} f(x) \, dx
\]

when \( y = O \) and

\[
\bar{f}(y) = \frac{1}{M_k(H_k)} \int_{C_k(H_k)} f(x) \frac{d\sigma}{|\nabla H_k(x)|_{\mathbb{R}^d}}
\]

when \( y = (k, H_k) \). The function \( v(y) \) satisfies the gluing condition

\[
0 = \text{Volume}(\mathcal{E}) \cdot \text{Av}(O) + \frac{1}{2} \sum_{k=1}^{r} p_k \cdot \lim_{H_k \to 0} \frac{\partial v}{\partial H_k}(k, H_k) .
\]

and \( v(\mathcal{Y}(x_O)) = 0 \). Such a function \( v(y) \) is unique.
Why $Y^\varepsilon_t \to Y_t$ weakly in $C_{[0,T]}(\Gamma)$? Heuristics.

- This proof of the convergence follows from the argument of Dolgopyat & Koralov, to appear in *Journal of AMS*. The basic idea can also be found in the classical monograph of Freidlin & Wentzell, Chapter 8.

- Situation is similar but actually simpler than the case of averaging for Hamiltonian flows on $\mathbb{T}^2$.

- For motion of $X^\varepsilon_t$ inside each of the $[U_k]$’s, the result is just a consequence of averaging principle, as we just did before in our example.

- For the motion of $X^\varepsilon_t$ in $\mathcal{E}$, the reason is that we have glued all points in $\mathcal{E}$ into one point $O$. 
Why $Y^\varepsilon_t \to Y_t$ weakly in $C_{[0,T]}(\Gamma)$? Heuristics.

- For fixed $\varepsilon > 0$ Lebesgue measure is invariant for the process $X^\varepsilon_t$ in $[G]$.
- When we do the projection $\mathcal{Y} : [G] \to \Gamma$, the limiting process $Y_t$, as a result, is expected to have an invariant measure $\mu$ on $\Gamma$, which is induced by Lebesgue measure of $X^\varepsilon_t$ on $[G]$.
- In particular, $\mu(\{O\}) = \text{Volume}(\mathcal{E})$.
- We have
  \[ \int_{\Gamma} Au(k, h) d\mu = 0. \]
- The above relation, when expanded, gives the gluing condition:
  \[ 0 = \text{Volume}(\mathcal{E}) \cdot \text{Av}(O) + \frac{1}{2} \sum_{k=1}^{r} p_k \cdot \lim_{H_k \to 0} \frac{\partial v}{\partial H_k}(k, H_k). \]
The proof: a short review of technicalities.

- The proof of convergence $\mathcal{Y}(X_t^\varepsilon) \to Y_t$ in $C_{[0,T]}(\Gamma)$ makes use of martingale problem techniques. I will omit the technicalities in this point here.

- We need to show that the process $X_t^\varepsilon$, as $\varepsilon$ is small, quickly tend to its invariant measure.

- This is not immediately obvious since as $\varepsilon$ is small the process $X_t^\varepsilon$ is close to a degenerate one near $\gamma$. In other words, we expect the process $X_t^\varepsilon$ to move slower and slower when it approaches $\gamma$.

- The key underlying reason that this is true is because of our assumption of "quadratic degeneracy":

$$\text{const}_1 \cdot \text{dist}^2(x, \gamma) \leq e_d(x) \cdot (a^{(0)}(x)e_d(x)) \leq \text{const}_2 \cdot \text{dist}^2(x, \gamma).$$
The proof: a short review of technicalities.

- We need the control of certain stopping times:

\[ \lim_{\varepsilon \downarrow 0} \sup_{x \in [\varepsilon]} E_x \sigma = 0, \]
\[ \lim_{\varepsilon \downarrow 0} \sup_{x \in \gamma} E_x \tau = 0, \]
\[ \lim_{\varepsilon \downarrow 0} \sup_{x \in \gamma} E_x \sigma = 0. \]

- These are estimated via the construction of barrier functions.
**Fig.** Details in the proof.
The proof: a short review of technicalities.

- Geometric construction: extension of the first integral $H_k$ to a neighborhood outside $\gamma$.
- "Global barrier": first introduce a Riemannian coordinate $(\varphi_1^k, \ldots, \varphi_{d-1}^k, H_k)$ near $\gamma$.
- "Barrier" function $u_k = u_k(H_k)$.
- Making use of some basic facts in Riemannian geometry it is possible to show that

\[
\left( \frac{1}{\varepsilon} L_0 + L_1 \right) u_k(x)
= \frac{1}{A(x)} \left[ \frac{\partial}{\partial H_k} \left( \left( \frac{K_1(x)}{\varepsilon} + K_2(x) \right) \frac{du_k}{dH_k}(H_k) \right) + K_3(x) \frac{du_k}{dH_k}(H_k) \right].
\]
The proof: a short review of technicalities.

- Making use of these barriers accompanied by some **dangerous** estimates our process $X_t^\varepsilon$ is able to freely travel in and out of $\gamma$.
- This is the main technical part of the work yet I will omit it in this talk.
The End.

- Thank you for your attention!