ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ

ΣΧΟΛΗ ΕΦΑΡΜΟΣΜΕΝΩΝ ΜΑΘΗΜΑΤΙΚΩΝ ΚΑΙ ΦΥΣΙΚΩΝ ΕΠΙΣΤΗΜΩΝ

ΜΕΤΑΠΤΥΧΙΑΚΗ ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ
ΔΠΜΣ ΕΦΑΡΜΟΣΜΕΝΕΣ ΜΑΘΗΜΑΤΙΚΕΣ ΕΠΙΣΤΗΜΕΣ

Martingale Problem for Diffusions in $\mathbb{R}^d$

Ιωάννα Βλάχου

Επιβλέπων Καθηγητής: Μιχαήλ Λουλάκης

Αθήνα, Ιούλιος 2023

Θα ήθελα να ευχαριστήσω τον επιβλέποντα καθηγητή μου, κύριο Μιχαήλ Λουλάκη για την πολύτιμη βοήθεια και καθοδήγηση που μου προσέφερε

Contents

0 Περίληψη

Οι διαδικασίες που μελετώνται σε αυτήν την εργασία είναι συνεχείς στοχαστικές μαρκοβιανές διαδικασίες συνεχούς χρόνου και συγκεκριμένα διαδικασίες οι οποίες προκύπτουν ως λύσεις στοχαστικών διαφορικών εξισώσεων. Για να μιλήσει κανείς για τέτοιες δαδικασίες πρέπει πρώτα να χτίσει ένα κατάλληλο πιθανοθεωρητικό χώρο και αυτό κανουμε στην εισαγωγή.
Αντί να δουλεύουμε σε έναν χώρο πιθανότητας με πολλές διαφορετικές διαδικασίες κάθε φορά, αυτό που κάνουμε είναι να θεωρήσουμε έναν μετρήσιμο χώρο $(\Omega, \mathcal{F})$ και μία οικογένεια μετρήσιμων τυχαίων μεταβλητών μέσα σε αυτόν $\{X_t\}_{t\geq 0}$ έτσι ώστε αλλάζοντας το μέτρο πιθανότητας στον $(\Omega, \mathcal{F})$, ουσιαστικά αλλάζουμε την κατανομή της διαδικασίας X. Επειδή θα περιοριστούμε μόνο σε συνεχείς d-διάστατες διαδικασίες, ο δειγματικός χώρος που παίρνουμε είναι ο χώρος των συνεχών συναρτήσεων από το $[0, \infty)$ στον $\mathbb{R}^d$ που συμβολίζουμε $C[0, \infty)^d$. Κατόπιν, θεωρούμε κάθε χρονική στιγμή την φυσική προβολή π_t που στέλνει κάθε συνεχή συνάρτηση ω στο $\omega(t)$. Ορίζοντας $X_t(\omega) = \omega(t)$ παίρνουμε μία συνεχή διαδικασία $\{X_t\}_t$ η οποία ονομάζεται κανονική. Συμβολίζουμε $\sigma(X_t : t \geq 0)$ τη σ-άλγεβρα που γεννιέται από την X και $\{\sigma(X_s : s \leq t)\}_{t\geq 0}$ την φυσική διήθησή της. Θεωρούμε $\mathbb{P}$ ένα μέτρο πιθανότητας. Επειδή θέλουμε διαδικασίες όπως το $\sup_{s\leq t} f(X_s)$ ή τυχαίες μεταβλητές όπως οι χρόνοι διακοπής να είναι μετρήσιμες και επειδή θέλουμε να συμπεριλάβουμε στα μετρήσιμα ενδεχόμενα σύνολα της μορής

$$A := \{\omega \in \Omega : X(\omega) \text{ έχει δεξιά παράγωγο στο } 0\}$$

θα θεωρήσουμε τις εξής εκδοχές των φυσικών σίγμα αλγεβρών που παράγει η κανονική διαδικασία. Συγκεκριμένα ορίζουμε $\mathcal{F}_t = \bigcap_{\epsilon>0} \bar{\mathcal{G}}_{t+\epsilon}$ όπου $\bar{\mathcal{G}}_s$ είναι η πλήρωση της $\sigma(X_r : r \leq s)$ ως προς $\mathbb{P}$ και $\mathcal{F} = \bigvee_{t\geq 0} \mathcal{F}_t$. Η διήθηση $\{\mathcal{F}_t\}_t$ λέγεται τότε ότι ικανοποιεί τις συνήθεις συνθήκες.

Ο χώρος των συνεχών συναρτήσεων μπορεί να γίνει πλήρης και διαχωρίσιμος έαν τον εφοδιάσουμε με κατάλληλη μετρική. Μία τέτοια μετρική είναι η

$$D(\omega, \omega') = \sum_{n=1}^{\infty} \frac{1}{2^n} \frac{\sup_{0\leq t\leq n} |\omega(t) - \omega'(t)|}{1 + \sup_{0\leq t\leq n} |\omega(t) - \omega'(t)}$$

Αυτό σημαίνει ότι τώρα μπορούμε να βρούμε και στον χώρο $\mathcal{P}(\Omega)$ μία μετρική που να τον κάνει πλήρη και διαχωρίσιμο και με αυτόν τον τρόπο μπορούμε να περιγράψουμε τα συμπαγή σύνολα ως ακολουθιακά συμπαγή. Ο σκοπός είναι να φτιάξουμε μία τοπολογία ώστε να μπορέσουμε να περιγράψουμε την σύγκλιση μέτρων πιθανότητας στον χώρο που ορίσαμε παραπάνω. Μία ακολουθία μέτρων πιθανότητας $\mathbb{P}_n$ συγκλίνει ασθενώς σε ένα μέτρο πιθανότητας $\mathbb{P}$ εάν $\mathbb{E}_{\mathbb{P}_n}[f] \underset{n}{\rightarrow} \mathbb{E}_{\mathbb{P}}[f]$ για κάθε συνεχή και φραγμένη συνάρτηση f του $\mathbb{R}^d$. Η μετρική του Prohorov εφοδιάζει τον χώρο $\mathcal{P}(\Omega)$ των μέτρων πιθανότητας με αυτή την τοπολογία και ταυτόχρονα τον κάνει πλήρη και διαχωρίσιμο. Ο παραπάνω ορισμός είναι συχνά δύσχρηστος και θα χρησιμοποιήσουμε, πολλές φορές, άλλα κριτήρια ώστε να συμπεράνουμε την ασθενή σύγκλιση. Το λήμμα Portmanteau είναι το πρώτο θεώρημα που δίνει ισοδύναμες συνθήκες ώστε μία ακολουθία μέτρων πιθανότητας να συγκλίνει ασθενώς.
Τα συμπαγή σύνολα στον (Ω, D) χαρακτηρίζονται από το θεώρημα Arzelà-Ascoli(Bourbaki) και ένα παρόμοιο θεώρημα στον $(\mathcal{P}(\Omega), d_P)$ δίνει ικανές και αναγκαίες συνθήκες ώστε ένα υποσύνολο $A \subset \mathcal{P}(\Omega)$ να είναι tight.

Ορισμός Μία οικογένεια μέτρων πιθανότητας Π στον $(\Omega, \mathcal{F})$ λέγεται tight εάν, για κάθε $\epsilon > 0$, υπάρχει συμπαγές σύνολο K ώστε $\inf_{\mathbb{P}\in\Pi} \mathbb{P}(K) > 1 - \epsilon$.

Το βασικότερο αποτέλεσμα που θα χρησιμοποιήσουμε εδώ, είναι το θεώρημα Prohorov που μεταφράζει το tightness μίας οικογένειας μέτρων πιθανότητας σε συμπάγεια της κλειστής της θήκης.

Ορισμός Ένα σύνολο A σε έναν μετρικό χώρο (X, d) λέγεται σχετικά συμπαγές εάν το $\bar{A}$ είναι συμπαγές.

Θεώρημα(Prohorov) Μία οικογένεια Borel μέτρων πιθανότητας στον Ω είναι tight εάν και μόνο εάν είναι σχετικά συμπαγής.

Το θεώρημα Prohorov είναι σε ισχύ σε κάθε πλήρη και διαχωρίσιμο μετρικό χώρο και θα το χρησιμοποιήοσουμε ως βασικό κριτήριο για να συμπεράνουμε τη σύγκλιση μέτρων. Για να το κάνουμε αυτό είναι αρκετό να δείξουμε μοναδικότητα των υπακολουθιακών ορίων(*Θεώρημα 2.6 Billingshley*).

Στη συνέχεια γίνεται μία σύντομη αναφορά στα αντίστοιχα θεωρήματα στον μεγαλύτερο χώρο των δεξιά συνεχών συναρτήσεων και της Shkorohod τοπολογίας. Αφού παρουσιαστεί η τοπολογική δομή του χώρου, γίνεται μία εισαγωγή στα θεωρήματα στοχαστικής ανάλυσης που θα χρησιμοποιήσουμε. Τα στοχαστικά ολοκληρώματα που εμφανίζονται παρακάτω είναι ολοκληρώματα ως προς συνεχή τετραγωνικά ολοκληρώσιμα martingales και για να έχουν νόημα πρέπει οι ως προς ολοκλήρωση διαδικασίες να πληρούν περεταίρω ιδιότητες μετρησιμότητας ή ολοκληρωσιμότητας.

Ορισμός Μία διαδικασία Y που ζει σε έναν χώρο πιθανότητας $(E, \mathcal{G}, \mathbb{P})$, ονομάζεται progressive εάν κάθε χρονική στιγμή t, η τυχαία μεταβλητή Y_t είναι μετρήσιμη ως προς τη σ-άλγεβρα $\mathbb{B}([0,t]) \times \mathcal{G}_t$.

Παρατήρηση Στον χώρο των συνεχών συναρτήσεων, η κανονική διαδικασία είναι πάντοτε progressive.

Το θεώρημα Lèvy λέει ότι η κίνηση Brown είναι το μοναδικό συνεχές martingale W με τετραγωνική κύμανστη $\langle W \rangle_t = t$, δηλαδή $W^2 - t$ είναι επίσης martingale.
Κάθε στοχαστικό ολοκλήρωμα $\int_0^t Y_s dW_s$, με την προϋπόθεση ότι η προσαρμοσμένη διαδικασία Y ικανοποιεί τη σχέση $\mathbb{E}[\int_0^t Y_s^2 ds] < \infty$ για κάθε t, είναι ένα συνεχές martingale. Το αντίστροφο είναι το θεώρημα αναπαράστασης των martingales.
Το τελευταίο θεώρημα στο εισαγωγικό μέρος, είναι ο μετασχηματισμός Girsanov που είναι το εξής αποτέλεσμα: Σε έναν χώρο πιθανότητας $(\Omega, \mathcal{F}, \mathbb{P})$, η κλάση των συνεχών martingales ως προς μία διήθηση που ικανοποιεί τις συνήθεις συνθήκες, παραμένει αναλλοίωτη μέσα από την αλλαγή μέτρου πιθανότητας με την προϋπόθεση ότι το νέο μέτρο είναι απόλυτα συνεχές με το $\mathbb{P}$. Η απόδειξη που επιλέγουμε είναι αυτή του *Cherny [2002]* που κάνει χρήση του θεωρήματος συμπάγειας του Prohorov.

Μία d-διάστατη στοχαστική διαδικασία $X_t = (X_t^{(1)}, ..., X_t^{(d)})$ σε έναν χώρο πιθανότητας $(\Omega, \mathcal{F}, \mathbb{P})$ είναι λύση στη στοχαστική διαφορική εξίσωση (SDE)

$$dX_t = b(X_t)dt + \sigma(X_t)dW_t$$

όπου W μία d-διάστατη κίνηση Brown, εάν ικανοποιεί την ολοκληρωτική εξίσωση

$$X_t = X_0 + \int_0^t b(X_s)ds + \int_0^t \sigma(X_s)dW_s \tag{0.1}$$

εάν επιπλέον η X είναι προσαρμοσμένη στη διήθηση που παράγει η κίνηση Brown W, τότε η λύση λέγεται ισχυρή.
Λέμε ότι η λύση X είναι μοναδική, εάν Y είναι μία άλλη διαδικασία που ικανοποιεί την εξίσωση (0.1) τότε οι δύο διαδικασίες X και Y είναι μη διακρινόμενες, δηλαδή $\mathbb{P}(X_t = Y_t, \ \forall t \geq 0) = 1$.

Ορισμός Μία στοχαστική διαδικασία X_t λέγεται ασθενής λύση στη SDE, εάν υπάρχει κάποια κίνηση Brown που ζεί στον ίδιο χώρο πιθανότητας $(\Omega, \mathcal{F}, \mathbb{P})$, τέτοια ώστε, το ζεύγος (X, W) να ικανοποιεί την ολοκληρωτική εξίσωση (0.1).

Ορισμός Η λύση στη SDE λέγεται ότι είναι ασθενώς μοναδική, εάν (Y, B) είναι μία άλλη ασθενής λύση της, τότε η B είναι κίνηση Brown και η Y έχει την ίδια οικογένεια κατανομών πεπερασμένης διάστασης με την X.

Κάθε ισχυρή λύση σε μία στοχαστική διαφορική εξίσωση είναι ασθενής λύση, το αντίστροφο όμως δεν ισχύει. Ένα παράδειγμα SDE που έχει ασθενή λύση αλλά δεν επιδέχεται ισχυρή είναι η εξίσωση $dX_t = sgn(X_t)dW_t$ όπου sgn είναι η συνάρτηση προσήμου.

Για τις λύσεις στοχαστικών διαφρικών εξισώσεων ισχύει ότι οι συνθέσεις τους με δύο φορές συνεχώς παραγωγίσιμες συναρτήσεις είναι επίσης λύσεις στοχαστικών διαφορικών εξισώσεων. Αυτό το αποτέλεσμα είναι γνωστό ως τύπος του Itô και διατυπώνεται ως εξής

$$f(X_t) = f(X_0) + \int_0^t \frac{1}{2} \sum_{i,j=1}^d \sigma_{ij}(X_s)\sigma_{ij}^T(X_t)\partial_{ij} f(X_s) + \sum_{i=1}^d b_i(X_s)\partial_i f(X_s) ds + \int_0^t \sigma(X_s)^T \nabla f(X_s) dW_s \tag{0.2}$$

Θεωρούμε τον γραμμικό διαφορικό τελεστή $\mathcal{L}$ που δρα πάνω σε C^2 συναρτησεις του $\mathbb{R}^d$ ως εξής

$$\mathcal{L}f(x) = \frac{1}{2} \sum_{i,j=1}^d a_{ij}(x)\partial_{ij} f(x) + \sum_{i=1}^d b_i(x)\partial_i f(x)$$

όπου $a = \sigma\sigma^T$ ονομάζεται συντελεστής διάχυσης και b συντελεστής τάσης. Τώρα ο τύπος του Itô παίρνει τη μορφή

$$f(X_t) = f(X_0) + \int_0^t \mathcal{L}f(X_s)ds + \int_0^t \sigma^T(X_s)\nabla f(X_s)dW_s$$

Το παρακάτω παράδειγμα (*Problem 5.4.4. Karatzas & Shreve [1991]*) δίνει την ιδέα πίσω από το πρόβλημα martingale. Θεωρείστε ότι W_t είναι μία μονοδιάστατη κίνηση Brown. Τότε η W είναι λύση στοχαστικής διαφορικής εξίσωσης με τετριμμένο τρόπο και ο τελεστής $\mathcal{L}$ όταν δρα πάνω σε μία $f \in C^2(\mathbb{R})$ γράφεται αντίστοιχα ως εξής

$$\mathcal{L}f(x) = \frac{1}{2} f''(x)$$

ενώ ο τύπος του Itô δίνει

$$f(W_t) = f(W_0) + \int_0^t \frac{1}{2} f''(W_s)ds + \int_0^t f'(W_s)dW_s$$

Το στοχαστικό ολολήρωμα είναι πεπερασμένο σχεδόν βεβαίως αφού η κίνηση Brown είναι συνεχής και συνεπώς φραγμένη σχεδόν βεβαίως από κάποιο συμπαγές υποσύνολο A του $\mathbb{R}$ σε συμπαγή διαστήματα χρόνου $[0, t]$ και έτσι και η συνεχής f' είναι φραγμένη στο A. Αυτό σημαίνει ότι η διαδικασία $\int_0^t f'(W_s)dW_s$ είναι ένα συνεχές local martingale.
Αντίστροφα, εάν μία στοχαστική διαδικασία W έχει την ιδιότητα ότι για κάθε $f \in C^2(\mathbb{R})$ η

$$f(W_t) - f(W_0) - \frac{1}{2} \int_0^t f''(W_s)ds$$

είναι ένα συνεχές local martingale τότε εάν επιλέξει κανείς πρώτα $f(x) = x$ και κατόπιν $f(x) = x^2$ βλέπει ότι οι διαδικασίες W και $W^2 - t$ είναι συνεχείς local martingales. Από το θεώρημα του Lèvy αυτό σημαίνει ότι η W είναι μία κίνηση Brown.
Το πρόβλημα martingale που θεμελειώθηκε από τους *Stroock & Varadhan* γενικεύει αυτόν τον martingale χαρακτηρισμό για d-διάστατες διαχύσεις δηλαδή για ασθενείς λύσεις στοχαστικών διαφορικών εξισώσεων.

Το πλαίσιο στο οποίο μελετάμε το πρόβλημα martingale είναι αυτό που ορίστηκε στην αρχή, ο χώρος των συνεχών συναρτήσεων με την συνήθη εκδοχή της Borel διήθησης που παράγει η κανονική διαδικασία.

Ορισμός Ένα μέτρο πιθανότητας $\mathbb{P}$ είναι λύση στο πρόβλημα martingale για τον τελεστή $\mathcal{L}$ με αρχική συνθήκη $x \in \mathbb{R}^d$ εάν

$$\mathbb{P}(X_0 = x) = 1$$

και η διαδικασία

$$f(X_t) - f(X_0) - \int_0^t \mathcal{L}f(X_s)ds$$

είναι martingale για κάθε $f \in C^2(\mathbb{R}^d)$.

Τα ερωτήματα που δημιουργούνται αφού δώσει κανείς αυτόν τον ορισμό είναι
1. Κάτω από ποιές συνθήκες (των συντελεστών του τελεστή) το πρόβλημα martingale έχει λύση;
2. Πότε η λύση του προβλήματος martingale είναι μοναδική;

Το πρόβλημα martingale για τον τελεστή $\mathcal{L}$ με αρχική συνθήκη x που συμβολίζουμε ως MP$(\mathcal{L}, \delta_x)$ λέγεται καλώς τεθημένο εάν επιδέχεται λύση η οποία είναι μοναδική.

Θα μπορούσαμε να θεωρήσουμε ως αρχική συνθήκη ένα οποιοδήποτε μέτρο στον $(\Omega, \mathcal{F})$. Αυτή είναι η πρόταση 1.31. Στην πρόταση 1.32. βλέπουμε ότι για να συμπεράνει κανείς μοναδικότητα της λύσης στο πρόβλημα martingale αρκεί να δείξει ότι δύο λύσεις έχουν κοινές οικογένειες κατανομών μίας διάστασης, δηλαδή ότι η κανονική διαδικασία κάθε χρονική στιγμή μέσω των δύο μέτρων λύσεων έχει την ίδια κατανομή.

Το τελευταίο αποτέλεσμα στο εισαγωγικό κεφάλαιο, είναι ένα κριτήριο των *Stroock & Varadhan* που συσχετίζει προβλήματα martingale με τις συμπαγείς οικογένειες μέτρων πιθανότητας στον χώρο των μέτρων $\mathcal{P}(\Omega)$. Υπάρχει επίσης ένα ανάλογο κριτήριο για διαδικασιές διακριτού χρόνου που θα χρησιμοποιήσουμε στο κεφάλαιο 4 όταν θα αντικαταστήσουμε τη κανονική μας διαδικασία με μία διακριτοποίησή της. Κάτω από τις υποθέσεις

1. Για κάθε $f \in C_0^\infty(\mathbb{R}^d)$ μη-αρνητική, η διαδικασία $f(X_t) + A_f t$ είναι submartingale για κάποια θετική σταθερά A_f
2. Δοθείσης μίας τέτοιας συνάρτησης f, η σταθερά A_f είναι η ίδια για όλες τις μεταθέσεις στο πεδίο ορισμού της f

έδωσαν μία εκτίμηση του ελάχιστου χρόνου για τον οποίον η οικογένεια $\{X_t\}$ παύει να είναι ισοσυνεχής, δηλαδή μία εκτίμηση για το modulus of continuity της X.

Στο κεφάλαιο 2 μελετάμε τη σχέση των λύεων του προβλήματος martingale με αυτές τις αντίστοιχης στοχαστικής διαφορικής εξίσωσης. Εν γέννει με τον τρόπο που ορίστηκε, το πρόβλημα martingale είναι ασθενέστερο από τις ασθενείς λύσεις της SDE, όμως στην πραγματικότητα η στοχαστική εξίσωση είναι καλώς τεθημένη εάν και μόνο εάν το πρόβλημα martingale είναι καλώς τεθημένο. Η απόδειξη γίνεται σε δύο βήματα και βασίζεται από την μία πλευρά στον τύπο του Itô και από την άλλη στο θεώρημα χαρακτηρισμού της κίνησης Brown του Lèvy. Στην αρχή παίρνουμε το αποτέλεσμα με την προϋπόθεση ότι ο τελεστής είναι ομοιόμορφα ελληπτικός και στη συνέχεια χρησιμοποιώντας το θεώρημα αναπαράστασης των martingale θα δούμε ότι ακόμα και για degenerate συντελεστές διάχυσης η ισοδυναμία ισχύει.

Ορισμός Ο τελεστής $\mathcal{L}$ λέγεται ομοιόμορφα ελλειπτικός εάν

$$\sup_i ||b_i(x)||_\infty < \infty$$

και υπάρχει θετικός αριθμός Λ τέτοιος ώστε

$$\Lambda|y|^2 \leq \sum_{i,j=1}^{d} y_i a_{ij}(x) y_j \leq \Lambda^{-1}|y|^2 \quad \text{για κάθε } x, y \in \mathbb{R}^d$$

Το ότι η ύπαρξη ασθενούς λύσης στη στοχαστική διαφορική εξίσωση $dX_t = b(X_t)dt + \sigma(X_t)dW_t$ συνεπάγει ύπραξη λύσης στο πρόβλημα martingale για τον τελεστή $\mathcal{L}$ που στέλνει μία $f \in C^2(\mathbb{R}^d)$ στη συνάρτηση $\mathcal{L}f$ με τύπο $\mathcal{L}f(x) = \frac{1}{2}\sum_{i,j=1}^d (\sigma\sigma^T)_{ij}(x)\partial_{ij}f(x) + \sum_{i=1}^d b_i(x)\partial_i f(x)$ προκύπτει πάλι όπως και στο παράδειγμα με την 1-διάστατη κίνηση Brown, από τον τύπο του Itô. Αντίστροφα, εάν υποθέσουμε ότι $\mathbb{P} \in MP(\mathcal{L}, \delta_x)$, τότε επιλέγοντας $f(x) = x_i$ και κατόπιν $f(x) = x_i x_j$ έπεται ότι για κάθε i, η διαδικασία

$$M^{(i)} = X_t^{(i)} - X_0^{(i)} - \int_0^t b_i(X_s)ds$$

είναι martingale διαδικασία με τετραγωνική κύμανση $\langle X^{(i)}, X^{(j)} \rangle_t = \int_0^t a_{ij}(X_s)ds$. Σε αυτό το σημείο ορίζει κανείς μία νέα διαδικασία ως εξής

$$W_t = \int_0^t \sigma^{-1}(X_s)dM_s^{id}$$

Εφόσο ο συμμετρικός πίνακας συναρτήσεων $a = \sigma\sigma^T$ είναι ομοιόμορφα ελληπτικός και άρα θετικά ορισμένος, η σ^{-1} είναι καλώς ορισμένη και συνεπώς η W είναι καλώς ορισμένη. Επιπλέον η διαδικασία $\int_0^t \sigma^{-1}(X_s)dM_s^{id}$ είναι ένα συνεχές τετραγωνικα ολοκληρώσιμο martingale αφού

$$\mathbb{E}\Big[{W_t^{(i)}}^2\Big] = \mathbb{E}\left[\int_0^t \sum_{n,m=1}^{d} (a^{-1}a)(X_s)ds\right] = t < \infty$$

Όμοια η τετραγωνική κύμανση προκύπτει $\mathbb{E}\Big[W_t^{(i)}W_t^{(j)}\Big] = t\delta_{ij}$. Από το θεώρημα Lèvy, η W είναι μία d-διάστατη κίνηση Brown και συνεπώς το ζεύγος (X, W) είναι λύση στην $dX_t = b(X_t)dt + \sigma(X_t)dW_t$.

Το πρόβλημα στην γενικότερη περίπτωση όπου ο συντελεστής διάχυσης είναι συμμετρικός θετικά ημιορισμένος είναι στο ότι δεν μπορούμε να ορίσουμε τη κίνηση Brown όπως κάναμε παραπάνω για την οποία η martingale διαδικασία

$X_t - X_0 - \int_0^t b(X_s)ds$ γράφεται ως στοχαστικό ολοκλήρωμα. Όμως επεκτείνοντας τον χώρο με κατάλληλο τρόπο θα μπορούμε να βρούμε μία νέα κίνηση Brown W που ζεί σε αυτόν τέτοια ώστε η M_t^{id} να μπορεί να γραφεί ως $\int_0^t \sigma(X_s)dW_s$.

Θεώρημα Έστω $\mathbb{P}$ μία λύση στο πρόβλημα martingale για τον τελεστή με φραγμένους μετρήσιμους συντελεστές $a = \sigma\sigma^T$ και b και αρχική συνθήκη x. Τότε υπάρχει μία d-διάστατη κίνηση Brown W που ζεί σε μία επέκταση $(E, \mathcal{G}, Q)$ του $(\Omega, \mathcal{F}, \mathbb{P})$ τέτοια, ώστε το ζεύγος (X, W) είναι ασθενής λύση στη SDE με συντελεστές b και σ που ξεκινάει από το σημείο x.
Η ιδέα της απόδειξης είναι η εξής. Από την υπόθεση και από το θεώρημα αναπαράστασης των martingales υπάρχει επέκταση $(E, \mathcal{G}, Q)$ του $(\Omega, \mathcal{F}, \mathbb{P})$ και επέκταση $\{\mathcal{G}_t\}_t$ της $\{\mathcal{F}_t\}_t$ καθώς και μία $\mathcal{G}_t$- προσαρμοσμένη κίνηση Brown B_t που ζεί σε αυτόν τέτοια ώστε

$$M_t = X_t - X_0 - \int_0^t b(X_s)ds = \int_0^t \xi_s dB_s$$

για κάποια $\mathcal{G}_t$- προσαρμοσμένη διαδικασία ξ για την οποία $\mathbb{E}[\int_0^t \xi_s^2 ds] < \infty$. Αρκεί να δείξουμε ότι

$$\int_0^t \xi_s dB_s = \int_0^t \sigma(X_s)dW_s$$

για κάποια κίνηση Brown W στον $(E, \mathcal{G}, Q)$.
Θεωρούμε τον $d \times d$ πίνακα απεικόνιση R που ορίζεται πάνω στο σύνολο

$$D = \{(\xi, \sigma) : \xi, \sigma \in M_{d\times d}(\mathbb{R}), \xi\xi^T = \sigma\sigma^T\}$$

με τις ιδιότητες ότι $\sigma R(\xi, \sigma) = \xi$ και $R(\xi, \sigma)R^T(\xi, \sigma) = I_{d\times d}$. Μία τέτοια απεικόνιση υπάρχει διότι οι πίνακες ξ, σ είναι συμμετρικοί και συνεπώς υπάρχουν ορθογώνιοι πίνακες U, V τέτοιοι ώστε $\sigma = U\Lambda_\sigma U^T$ και $\xi = V\Lambda_\xi V^T$. Τότε ο πίνακας $R = U^{-1}V$ πληρεί τις παραπάνω ιδιότητες. Στο σημείο αυτό παρατηρούμε ότι μία τέτοια απεικόνιση είναι Borel μετρήσιμη (Remark 2.3.). Έτσι εάν ορίσουμε

$$W_t = \int_0^t R^T(\xi_s, \sigma(X_s))dB_s$$

έχουμε μία καλώς ορισμένη $\mathcal{G}_t$-progressive συνεχή διαδικασία η οποία είναι martingale αφού

$$\mathbb{E}_Q\left[W_t^{(i)^2}\right] = \mathbb{E}_Q\left[\sum_{k,l=1}^{d} \int_0^t R_{ik}^T R_{il}(\xi_s, \sigma(X_s))ds\right] = t < \infty$$

Τώρα μπορούμε να χρησιμοποιήσουμε τον χαρακτηρισμό του Lèvy για την κίνηση Brown και όπως στην περίπτωση του ομοιόμορφα ελλειπτικού τελεστή έπεται ότι $dX_t = b(X_t)dt + \sigma(X_t)dW_t$

Στο τρίτο κεφάλαιο ασχολούμαστε με την ύπαρξη και μοναδικότητα των λύσεων του προβλήματος. Στο πρώτο θεώρημα βλέπουμε ότι η μοναδικότητα σχετίζεται με την ύπαρξη λύσεων σε ένα άλλο πρόβλημα · στο πρόβλημα Cauchy. Στην περίπτωση όπου το πρόβλημα martingale έχει μοναδική λύση, η διαδικασία είναι ισχυρά Μαρκοβιανή. Στη συνέχεια θα μελετήσουμε την ύπαρξη λύσεων στο MP ως εφαρμογή της σύγκλισης μέτρων πιθανότητας με βάση το θεώρημα 7.3. του Billingshley ή το κριτήριο των Stroock & Varadhan.

Θεώρημα Η ακολουθία μέτρων πιθανότητας $\{\mathbb{P}^n\}_n$ είναι tight στον $\mathcal{P}(\Omega)$ εάν και μόνο εάν ισχύουν τα παρακάτω
(i) για κάθε $\epsilon > 0$ υπάρχει $\zeta > 0$ τέτοιο ώστε $\mathbb{P}^n(|X_0| > \zeta) \leq \epsilon$ για κάθε $n \geq n_0$ για κάποιο θετικό ακέραιο n_0.
(ii) για κάθε $\epsilon > 0$, ισχύει ότι $\lim_{\delta\to 0}\limsup_n \mathbb{P}^n(w(\delta) \geq \epsilon) = 0$ όπου $w(\delta) = \sup\limits_{|t-s|<\delta} |X_t - X_s|$

Μία συνέπεια στην περίπτωση όπου το MP είναι καλά τεθημένο είναι ότι εάν μπορέσουμε να προσεγγίσουμε τον τελεστή με κατάλληλο τρόπο μπορούμε να συμπεράνουμε την ασθενή σύγκλιση διαχύσεων στην μοναδική λύση του MP (*Theorem 11.1.4. Stroock & Varadhan[1979]*).

Θεώρημα Υποθέτουμε ότι το MP$(\mathcal{L}, \delta_x)$ είναι καλά τεθημένο, για κάθε $x \in \mathbb{R}^d$ και ότι ο τελεστής $\mathcal{L}$ έχει συνεχείς ομοιόμορφα φραγμένους συντελεστές τάσης b και διάχυσης a με $a = \sigma\sigma^T$. Υποθέτουμε ακόμη ότι υπάρχουν ακολουθίες b_n και a_n τέτοιες ώστε

$$\sup_n \sup_{|x|\leq R} ||a_n(x)|| + |b_n(x)| < \infty$$

και

$$\lim_n \int_0^T \sup_{|x|\leq R} (||a_n(x) - a(x)|| + |b_n(x) - b(x)|)ds = 0$$

για κάθε $T > 0$ και $R > 0$. Εάν το MP$(\mathcal{L}_n, \delta_{x_n})$ έχει λύση με $\mathcal{L}_n$ να είναι ο τελεστής με συντελεστές a_n και b_n και $x_n \to x$ τότε $\mathbb{P}^n \Rightarrow \mathbb{P}$ όπου $\mathbb{P}$ είναι η μοναδική λύση του MP$(\mathcal{L}, \delta_x)$.

Απόδειξη : Θα δείξουμε ότι η ακολουθία $\mathbb{P}^n$ είναι σχετικά συμπαγής και ότι κάθε υπακολουθιακό όριό της Q είναι η μοναδική λύση $\mathbb{P}$ στο πρόβλημα martingale $MP(\mathcal{L}, \delta_x)$, μετά από το θεώρημα 2.6 του Billinshley έπεται το ζητούμενο.

Από το θεώρημα Prohorov αρκεί να δείξουμε ότι η ακολουθία $\mathbb{P}^n$ είναι tight. Παρατηρούμε ότι ισχύουν οι συνθήκες του θεωρήματος 7.3 του Billingshley.

(i) $\mathbb{P}^n(|X_0| > \sup_n |x_n| + 1) = 0$

(ii) Θεωρούμε Λ_a, Λ_b να είναι τα φράγματα των a, b αντίστοιχα. Έστω $\epsilon > 2\Lambda_b(t-s)$.

$$\{\sup_{s\le t\le s+\delta} |X_t - X_s| \ge \epsilon\} \subset \{\sup_{s\le t\le s+\delta} |X_t - X_s - \int_s^t b_n(X_r)dr| + \Lambda_b(t-s) \ge \epsilon\}$$

δηλαδή

$$\{\sup_{s\le t\le s+\delta} |X_t - X_s| \ge \epsilon\} \subset \{\sup_{t\le s+\delta} |M_t^s| + \Lambda_b(t-s) \ge \epsilon\} \subset \{\sup_{t\le s+\delta} |M_t^s| \ge \epsilon/2\}$$

όπου $M_t^s := X_{t+s} - X_s - \int_s^{t+s} b_n(X_r)dr$. Επειδή η a είναι φραγμένη έχουμε ότι η τετραγωνική κύμανση της διαδικασίας M_t^s, $\langle M_\cdot^s\rangle_t = \int_s^t a(X_r)dr$ φράσεται από το $\Lambda_a(t-s)$. Χρησιμοποιώντας το σχόλιο 3.2. παίρνουμε ότι

$$\mathbb{P}^n(\sup_{s\le t\le s+\delta} |X_t - X_s| \ge \epsilon) \le \mathbb{P}^n(\sup_{t\le s+\delta} |M_t^s| \ge \epsilon/2) =$$

$$\mathbb{P}^n(\sup_{t\le s+\delta} |M_t^s| \ge \epsilon/2, \langle M_\cdot^s\rangle_t \le \Lambda_a(t-s)) \le e^{\frac{-\epsilon^2}{8c(t-s)}}$$

όπου η σταθερά c δεν εξαρτάται από το n. Στέλνουμε το $\delta \to 0$ και έχουμε

$$\lim_{\delta\to 0} \limsup_n \mathbb{P}^n(w(\delta) \ge \epsilon) = 0$$

για κάθε $\epsilon > 0$.

Αυτό που μένει να δείξουμε είναι ότι οποιοδήποτε το υπακολουθιακό όριο λύνει το πρόβλημα martingale MP$(\mathcal{L}, \delta_x)$. $Q(X_0 = x) = 1$ αφού $x_{n_k} \to x$ και $\mathbb{E}_{\mathbb{P}^{n_k}}[h(X_0)] \to \mathbb{E}_Q[h(X_0)]$ για κάθε συνεχή και φραγμένη h. Για την ιδιότητα martingale αρχικά δείχνουμε την ιδιότητα για φραγμένες συναρτήσεις $f \in C^2(\mathbb{R}^d)$ με φραγμένες πρώτες και δεύτερες μερικές παραγώγους χρησιμοποιώντας το Λήμμα 3.5. Αν η συνάρτηση f δεν είναι απαραίτητα φραγμένη θεωρούμε f^R τη συνεχή συνάρτηση που συμφωνεί με την f στη μπάλα $B(0,R)$, τότε η σταματημένη test διαδικασία θα είναι martingale από την προηγούμενη περίπτωση. Τέλος στέλνουμε το $R \to \infty$.

Η πρώτη εφαρμογή αυτής της σύγκλισης είναι η ύπαρξη λύσεων σε προβλήματα martingale για τελεστές $\mathcal{L}$ με συνεχείς συντελεστές αφού μπορούμε να

βρούμε C^2 συναρτήσεις που συγκλίνουν ομοιόμορφα στα συμπαγή υποσύνολα του $\mathbb{R}^d$ σε αυτούς τους συνεχείς συντελεστές του τελεστή. Οι αντίστοιχες ασθενείς λύσεις SDE έχουν Lipschitz συνεχείς συντελεστες και συνεπως είναι μοναδικές και η οικογένεια λύσεων προκύπτει σχετικά συμπαγής όπως παραπάνω. Το υπακολουθιακό όριο είναι αυτό που λύνει το πρόβλημα martingale για τον $\mathcal{L}$.
Στην περίπτωση όπου ο τελεστής $\mathcal{L}$ είναι ομοιόμορφα ελλειπτικός, το πρόβλημα επιδέχεται λυση για αυθαίρετους Borel μετρήσιμους φραγμένους συντελεστές a και b. Για να το δείξουμε αυτό αρχικά δείχνουμε ότι κάτω από τη συνθήκη ελλειπτικότητας μπορούμε να απαλέιψουμε πλήρως τον συντελεστή τάσης και να δείξουμε ισοδύναμα ότι το πρόβλημα martingale για τον νέο τελεστή $\mathcal{L}' = \frac{1}{2}\sum_{i,j=1}^{d} a_{ij}\partial_{ij}$ έχει λύση. Αυτό γίνεται με τον μετασχηματισμό Girsanov. Μετά την απαλοιφή του b η ύπαρξη έρχεται από ένα υπακολουθιακό όριο λύσεων MP για C^2 τελεστές. Αρχικά μπορούμε να προσεγγίσουμε το συνελεστή διάχυσης a από συνεχείς συναρτήσεις (από το θεώρημα Lusin στη δεύτερη μορφή και το θεώρημα επέκτασης του Tietze) και επειδή ο $C^2(\mathbb{R}^d)$ είναι πυκνός στον C μπορούμε να βρούμε μία ακολουθία $\{a_n\}_n \subset C^2(\mathbb{R}^d)$ που συγκλίνει στην a σχεδόν παντού. Όπως και στην προηγούμενη εφαρμογή, τα αντίστοιχα προβλήματα martingale $MP(\mathcal{L}_n, \delta_x)$ είναι καλά τεθημένα και η ακολουθία των λύσεων $\mathbb{P}_n$ έχει υπακολουθία $\mathbb{P}^{n_k}$ που συγκλίνει σε κάποιο μέτρο πιθανότητας $\mathbb{P} \in \mathcal{P}(\Omega)$. Επειδή οι συντελεστές είναι φραγμένοι μπορούμε να θεωρήσουμε μία υπακολουθία της a_{n_k}, $a_{n_{k_m}}$ η οποία συγκλίνει στην a στον $\mathbb{L}_p$ για κάποιο $p \geq 1$. Αρκεί να δείξουμε τότε ότι $\lim_m \mathbb{E}_{\mathbb{P}^{n_{k_m}}}[\{M_t^f - M_s^f\}\mathbb{1}_A] = \mathbb{E}_{\mathbb{P}}[\{M_t^f - M_s^f\}\mathbb{1}_A]$ για κάθε $A \in \mathcal{F}_s$. Θα χρειαστούμε το παρακάτω θεώρημα (*Exercise 7.3.2. Stroock & Varadhan [1979]*).

Θεώρημα (Alexandrov's estimate) Για ομοιόμορφα ελλειπτικούς τελεστές και για κάθε $p > d$, $t > 0$, $R > 0$ και $f \in C_0^\infty(\mathbb{R}^d)$ με φορέα εντός της $B(0, R)$ υπάρχει σταθερά C τέτοια ώστε

$$\left|\mathbb{E}_{\mathbb{P}}\left[\int_0^t f(X_s)ds\right]\right| \leq C||f||_{\mathbb{L}_p}$$

με τη σταθερά C να εξαρτάται από το ελλειπτικό φράγμα Λ και τα p, t, R.

Με το ίδιο επιχείρημα όπως πριν και τους ίδιους χρόνους διακοπής τ_R αρκεί να δείξουμε ότι

$$\mathbb{E}_{\mathbb{P}^{n_{k_m}}}\left[\int_0^t \mathcal{L}_{n_{k_m}} f(X_r)dr\right] \underset{n}{\rightarrow} \mathbb{E}_{\mathbb{P}}\left[\int_0^t \mathcal{L} f(X_r)dr\right]$$

για συναρτήσεις f φραγμένες με φραγμένες πρώτες και δεύτερες παραγώγους. Σε αυτό το σημείο χρησιμοποιούμε το προηγούμενο θεώρημα.

Μία ενδιαφέρουσα εφαρμογή ως συνέπεια της μοναδικότητας λύσεων του προβλήματος martingale είναι ότι η ασθενής λύση στη στοχαστική διαφορική εξίσωση είναι όριο Μαρκοβιανών αλυσίδων. Στο κεφάλαιο 4 όπου θα μελετήσουμε αυτή τη σύγκλιση, θα θεωρήσουμε μία διακριτοποίηση της κανονικής μας διαδικασίας η οποία κάτω από ένα μέτρο πιθανότητας θα είναι μία Μαρκοβιανή αλυσίδα που ξεκινάει από ένα σημείο $x \in \mathbb{R}^d$ με κάποια συνάρτηση πιθανοτήτων μετάβασης. Τότε θα δούμε ότι εάν το πρόβλημα martingale είναι καλά τεθημένο και υπό προϋποθέσεις των συντελεστών και της συνάρτησης πιθανοτήτων μετάβασης, για κάθε αρχική κατάσταση $x \in \mathbb{R}^d$ οι κατανομές των Μαρκοβιανών αλυσίδων θα συγκλίνουν ασθενώς στη μοναδική λύση του MP (*Chpater 11.2 Stroock & Varadhan [1979]*).
Η διακριτοποίηση γίνεται ως εξής: Έστω $h > 0$, $x \in \mathbb{R}^d$ και $\Pi_h(x, \cdot)$ μία συνάρτηση πιθανοτήτων μετάβασης. Θεωρούμε ένα μέτρο πιθανότητας $\mathbb{P}_h^x$ στον $C(0,\infty]^d$ με τις ιδιότητες :
(i) $\mathbb{P}_h^x(X_0 = x) = 1$

(ii) $\mathbb{P}_h^x\bigg(X_t = \frac{(k+1)h-t}{h}X_{kh} + \frac{t-kh}{h}X_{(k+1)h}, \ \ kh \leq t < (k+1)h\bigg) = 1$

(iii) Για κάθε $\Gamma \in \mathbb{B}(\mathbb{R}^d)$, $\mathbb{P}_h^x(X_{(k+1)h\in\Gamma}|\mathcal{F}_{kh}) = \Pi_h(X_{kh}, \Gamma)$

Η $\{X_{kh}\}_{k\in\mathbb{N}}$ είναι μία αλυσίδα στον χώρο των πραγματικών ακολουθιών και για κάθε $\omega \in \Omega$ η τιμή $X_{kh}(\omega) = \omega(kh)$ είναι ο kh-όρος της ακολουθίας ω_n. Με την τρίτη ιδιότητα, ορίσαμε η X_{kh} να είναι Μαρκοβιανή αλυσίδα με συνάρτηση πιθανοτήτων μετάβασης $\Pi_h(x, \cdot)$ ως προς το μέτρο πιθανότητας $\mathbb{P}_h^x$.

Μετά την διακριτοποίηση της διαδικασίας χρειαζόμαστε τον ορισμό του προβλήματος martingale στη διακριτή περίπτωση.
Ορισμός Στον χώρο των πραγματικών ακολουθιών, ένα μέτρο πιθανότητας $\mathbb{P}$ λύνει το πρόβλημα martingale για τον τελεστή A_n με αρχική συνθήκη $x \in \mathbb{R}^d$ εάν $\mathbb{P}(X_0 = x) = 1$ και η διαδικασία

$$f(X_n) - f(X_0) - \sum_{i=1}^{n-1} A_i f(X_i)$$

είναι martingale διαδικασία για κάθε συνεχή και φραγμένη συνάρτηση f του $\mathbb{R}^d$. Εδώ ο τελεστής A_n για f στον $C_b(\mathbb{R}^d)$ δρα ως εξής

$$A_n f(x) = \int \{f(y) - f(x)\}\Pi_n(x, dy)$$

Η πρώτη πρόταση (*Exercise 6.7.1. Stroock & Varadhan [1979]*) είναι ένα διακριτό ανάλογο του θεωρήματος ισοδυναμίας του κεφαλαίου 2. Συγκεκριμένα το διακριτό πρόβλημα martingale $MP(A_n, \delta_x)$ έχει μοναδική λύση εάν και μόνο εάν η X_n είναι μία Μαρκοβιανή αλυσίδα με συνάρτηση πιθανοτήτων μετάβασης $\Pi_n(x, \cdot)$.

Υποθέστε ότι οι $a : \mathbb{R}^d \to M_{d\times d}(\mathbb{R})$ και $b : \mathbb{R}^d \to \mathbb{R}^d$ είναι ομοιόμορφα φραγμένες συνεχείς συναρτήσεις και ότι ο a είναι θετικά ημιορισμένος. Για κάθε $h > 0$ και $\epsilon > 0$ ορίζουμε συναρτήσεις a_h b_h και Δ_h^ϵ

$$a_h(x) := \int_{|x-y|\le 1} (y_i - x_i)(y_j - x_j)\Pi_h(x, dy)$$

$$b_i(x) := \int_{|x-y|\le 1} (y_i - x_i)\Pi_h(x, dy)$$

$$\Delta_h^\epsilon(x) := \frac{\Pi_h(x, B(x,\epsilon)^c)}{h}$$

Υποθέσεις

(i) Οι συναρτήσεις a_h και b_h είναι ομοιόμορφα φραγμένες.
(ii) $a_h \to a$ $b_h \to b$ ομοιόμορφα στα συμπαγή υποσύνολα του $\mathbb{R}^d$ όταν $h \to 0$.
(iii) $\lim_{h\to 0} \sup_{x\in\mathbb{R}^d} \Delta_h^\epsilon(x) = 0$ για κάθε $\epsilon > 0$.

Θεώρημα Έστω ότι το πρόβλημα martingale για τον τελεστή $\mathcal{L}$ με συντελεστές a και b είναι καλά τεθημένο. Έστω ακόμη ότι ισχύουν οι παραπάνω υποθέσεις. Εάν $\mathbb{P}^{x_0}$ είναι η μοναδική λύση του $MP(\mathcal{L}, \delta_{x_0})$ και εάν $\mathbb{P}_h^{x_0}$ είναι η μοναδική λύση του $MP(A_h, \delta_{x_0})$ με συνάρτηση πιθανοτήτων μετάβασης $\Pi_h(x_0, \cdot)$, τότε η οικογένεια $\mathbb{P}_h^{x_0}$ συγκλίνει ασθενώς στο $\mathbb{P}^{x_0}$ καθώς $h \to 0$.

1 Martingale Problem

1.1 Introduction

We start by introducing the probability space and the processes we are going to study. We restrict our attention only to continuous stochastic processes taking values in $\mathbb{R}^d$ in continuous time. We define the sampling space $\Omega := C([0,\infty])^d$ to be the space of continuous functions from $[0,\infty)$ to $\mathbb{R}^d$. This space is Polish and can be equipped with a uniform convergence on compacts topology. A metric that makes Ω complete and is compatible with this topology is

$$D(\omega,\omega') = \sum_{n=1}^{\infty} \frac{1}{2^n} \frac{\sup_{0\le t\le n} |\omega(t)-\omega'(t)|}{1+\sup_{0\le t\le n} |\omega(t)-\omega'(t)}$$

Now consider the Borel σ-algebra generated by the open sets of this topology. We are going to deal with probability measures on $(\Omega, \mathbb{B}(\Omega))$, we denote $\mathcal{P}(\Omega)$ the space of all such probability measures. $\mathcal{P}(\Omega)$ is Polish, since Ω is Polish. We will define on $\mathcal{P}(\Omega)$ a weak topology, that is a topology generated by sets

$$\{\mathbb{P} \in \mathcal{P}(\Omega) : |\int_{\Omega} f d\mathbb{P} - x| < \epsilon\}$$

Prohorov's metric defined as

$$d(\mu,\nu) := \inf\{\epsilon > 0 : \mu(F) \le \nu(F^{\epsilon}) + \epsilon, \forall F \text{ closed in } \Omega\}$$

where

$$F^{\epsilon} = \{\omega \in \Omega : D(\omega, F) < \epsilon\}$$

equips $\mathcal{P}(\Omega)$ with the weak topology and additionally make it complete. Now convergence may be defined in sense of convergence of sequences and compactness in sense of sequential compactness. For sets that are not necessarily closed we use the notion of relatively compactness.

Definition 1.1. A sequence of probability measures $\mathbb{P}_n$ in $\mathcal{P}(\Omega)$ is said to converge to a probability measure $\mathbb{P}$ in $\mathcal{P}(\Omega)$ and denote $\mathbb{P}_n \Rightarrow \mathbb{P}$, if

$$\int_{\Omega} f d\mathbb{P}_n \to \int_{\Omega} f d\mathbb{P} \ \ \forall f \in C_b(\mathbb{R}^d)$$

The first lemma gives equivalent conditions for convergence of probability measures (*Theorem 2.1. Billingshley [1999]*).

Theorem 1.2. (Lemma Portmanteau) Let
The following are equivalent
(i) $\mathbb{P}_n \Rightarrow \mathbb{P}$
(ii) $\int f d\mathbb{P}_n \to \int f d\mathbb{P}$, for every f bounded uniformly continuous function of $\mathbb{R}^d$
(iii) $\limsup_n \mathbb{P}_n(F) \leq \mathbb{P}(F)$, for every closed F in Ω
(iv) $\liminf_n \mathbb{P}_n(G) \geq \mathbb{P}(G)$, for every open G in Ω
(v) $\lim_n \mathbb{P}_n(A) = \mathbb{P}(A)$, for every Borel set A such that $\mathbb{P}(\partial A) = 0$

Proof: (ii) $\Longrightarrow$ (iii) Suppose $\int f(x)\mathbb{P}_n(dx) \to \int f(x)\mathbb{P}(dx)$ for every bounded uniformly continuous function f and let F closed set in Ω and $\epsilon > 0$. We consider the function $f(x) = \max\{0, 1 - \frac{d(x,F^\epsilon)}{\epsilon}\}$ which is a function both bounded and uniformly continuous and satisfies the following inequality

$$\mathbb{1}_F \leq f(x) \leq \mathbb{1}_{F^\epsilon}$$

consequently

$$\limsup_n \mathbb{P}_n(F) \leq \limsup_n \int f(x)\mathbb{P}_n(dx) = \int f(x)\mathbb{P}(dx) \leq \int \mathbb{1}_{F^\epsilon}\mathbb{P}(dx) = \mathbb{P}(F^\epsilon)$$

now let $\epsilon \to 0$, then $\limsup_n \mathbb{P}_n(F) \leq \mathbb{P}(F^0) = \mathbb{P}(F)$. The last equality is due to the fact that F is closed.

(iii) $\Longleftrightarrow$ (iv) We have $\liminf_n \mathbb{P}_n(F^c) \geq \mathbb{P}(F^c)$ for every closed set F in Ω.

(iv) $\Longrightarrow$ (v) Let $A \in \mathbb{B}(\Omega)$ such that $\mathbb{P}(\partial A) = 0$, then $\mathbb{P}(\bar{A}) = \mathbb{P}(A^\circ)$ and therefore

$$\mathbb{P}(\bar{A}) \geq \limsup_n \mathbb{P}_n(\bar{A}) \geq \liminf_n \mathbb{P}_n(A^\circ) = \mathbb{P}(A^\circ) = \mathbb{P}(\bar{A})$$

(v) $\Longrightarrow$ (i) Now consider a bounded continuous function f, then it is enough to show

$$\int \frac{f}{M} d\mathbb{P}_n \underset{n}{\to} \int \frac{f}{M} d\mathbb{P}$$

where M is a bound of f, thus we can assume that $0 < f < 1$. We get

$$\int f(x)\mathbb{P}(dx) = \int_0^\infty \mathbb{P}(f > t)dt = \int_0^1 \mathbb{P}(f > t)dt$$

and similarly

$$\int f(x)\mathbb{P}_n(dx) = \int_0^1 \mathbb{P}_n(f > t)dt$$

Because f is continuous we have $\partial\{f > t\} \subset \{f = t\}$ and $\mathbb{P}(f = t) > 0$ only for at most countable many t's, otherwise $\mathbb{P}(\Omega) = \infty$ which is untrue. Therefore $\mathbb{P}(\partial\{f > t\}) = 0$ except for at most countable many t's. Using the assumption in (v) we conclude by dominating convergence theorem

$$\lim_n \int_0^1 \mathbb{P}_n(f > t)dt = \int_0^1 \mathbb{P}(f > t)dt$$

implying

$$\lim_n \int f d\mathbb{P}_n = \int f d\mathbb{P}$$

Note 1.3. The same proof with implication $(ii) \implies (iii)$ gives also uniqueness of weak convergence limits.

Definition 1.4. Let (Y, ρ) is a metric space. A subset $A \subset Y$ is called relatively compact if every sequence of elements of A has convergent subsequence, its limit point may not lie in A.

Arzelá -Ascoli gave necessary and sufficient conditions for relatively compactness in $C[0,\infty)^d$, in section 3.3. we will see similar conditions for characterization of tightness in $\mathcal{P}(C[0,\infty)^d)$ *(Theorems 7.2-7.3 Billingshley[1999])*.

Theorem 1.5. (Arzelá - Ascoli in $C[0,\infty)^d$) A set A in $C[0,\infty)^d$ is relatively compact if and only if

$$\sup_{\omega \in A} |\omega(0)| < \infty$$

and

$$\lim_{\delta \to 0} \sup_{\omega \in A} w_\omega(\delta) = 0$$

where

$$w_\omega(\delta) := \sup_{|t-s|<\delta} |\omega(t) - \omega(s)|$$

is called the modulus of continuity of ω.

Theorem 1.2. is a translation of *Theorem X.5.2. N.Bourbaki [1966]* for real-valued continuous functions on $[0,\infty)$.

A theorem that gives description of relatively compact sets in $\mathbb{P}(\Omega)$ is Prohorov's Theorem, but first we need the definition of tightness.

Definition 1.6. A family probability measures Π on $(C[0,\infty)^d], \mathbb{B}(C[0,\infty)^d))$ is called tight if for each $\epsilon > 0$, there exists a compact set K such that

$$\inf_{\mathbb{P}\in\Pi} \mathbb{P}(K) > 1 - \epsilon$$

Remark 1.7. Every probability measure in $\mathcal{P}(C[0,\infty)^d])$ is tight.

Proof: Let $\mathbb{P}$ be a probability measure in $\mathcal{P}(\Omega)$ and let D be dense in Ω. We consider the collection $\mathcal{C} := \{B(x_k, \frac{1}{n}) : x_k \in D, n, k \in \mathbb{N}\}$. Then, $\mathcal{C}$ forms a base for the topology in Ω and therefore we can cover any open set by countable union of elements of $\mathcal{C}$, in particular $\Omega = \bigcup_{k=1}^{\infty} \bigcup_{n=1}^{\infty} B(x_k, \frac{1}{n})$. Let $\epsilon > 0$, choose m such that $\mathbb{P}(\bigcup_{n=1}^{m} B(x_{k_n}, \frac{1}{n})) > 1 - \frac{\epsilon}{2^m}$ and define

$$A := \bigcap_{m=1}^{\infty} \bigcup_{n=1}^{m} B(x_{k_n}, \frac{1}{n})$$

then $\bar{A}$ is compact. In particular $\bar{A}$ is bounded by $\bigcup_{n=1}^{m} B(x_{k_n}, 2)$. Moreover

$$\mathbb{P}(\bar{A}^c) \leq \mathbb{P}(A^c) = \sum_{m=1}^{\infty} \mathbb{P}((\bigcup_{n=1}^{m} B(x_{k_n}, \frac{1}{n}))^c) < \sum_{m=1}^{\infty} \frac{\epsilon}{2^m} = \epsilon$$

therefore $\mathbb{P}$ is tight.

Theorem 1.8. (Prohorov) Suppose Π is a family of Borel probability measures in Ω. The following are equivalent.

(i) Π is relatively compact in $\mathcal{P}(\Omega)$.

(ii) Π is tight

Proof: (i) $\implies$ (ii). By Remark 1.3. we know that every probability measure in $\mathcal{P}(\Omega)$ is tight. This time we have to find suitable compact set that works for all members of the family Π.
Claim: Consider an open covering $\{U_i\}_{i=1}^{\infty}$ of Ω. Because Π is relatively compact, for each $\epsilon > 0$, there is a subcovering $\{U_i\}_{i=1}^{k}$ such that $\mathbb{P}(\bigcup_{i=1}^{k} U_i) > 1 - \epsilon$.

Proof (Claim): Suppose $\epsilon > 0$ such that $\forall k \in \mathbb{N}$, there is $\mathbb{P}^k \in \Pi$ such that $\mathbb{P}^k(\bigcup_{i=1}^{k} U_i) \leq 1 - \epsilon$. Since the family Π is relatively compact, there is a subsequence $\mathbb{P}_j^k$ converging weakly to a probability measure $\mu \in \bar{\Pi}$. By Portmanteau's lemma we get

$$\mu(\bigcup_{i=1}^{k} U_i) \leq \liminf_j \mathbb{P}_j^k(\bigcup_{i=1}^{k} U_i) \leq \liminf_j \mathbb{P}_j^k(\bigcup_{i=1}^{k_j} U_i) \leq 1 - \epsilon$$

but since, $\Omega = \bigcup_{i=1}^{k} U_i$ then it must $\mu(\bigcup_{i=1}^{k}) \underset{k}{\rightarrow} \mu(\Omega) = 1$ which is untrue.

So the claim is valid. Let $\epsilon > 0$. Suppose D is dense in Ω, $\forall m \geq 1$ consider again Ω is covering by $\bigcup_{i=1}^{\infty} B(x_i, \frac{1}{m})$ where $x_i \in D$. By claim we know that there is k_m such that $\mathbb{P}(\bigcup_{i=1}^{k_m} B(x_i, \frac{1}{m})) > 1 - \frac{\epsilon}{2^m}$, for all $\mathbb{P} \in \Pi$. Similarly with Remark 1.3. the set $\bar{A}$ is compact satisfying $\mathbb{P}(\bar{A}) > 1 - \epsilon$ for every $\mathbb{P} \in \Pi$.

(ii) $\Longrightarrow$ (i). Suppose Π is tight. We want to show that for each sequence of measure in Π, there is a convergent subsequence. We will use two lemmas from functional analysis.

Lemma 1. If (X, d) is compact metric space, then $(\mathcal{P}(X), d_P)$ is compact metric space (d_P is the Prohorov metric related to d).

Lemma 2. If (X, d) is seperable metric space, it is homeomorphic to a compact metric space.

First we see that if Π is tight, then $\bar{\Pi}$ is also tight. Indeed, let $\epsilon > 0$ and K compact such that $\mathbb{P}(K) > 1 - \epsilon$ for each $\mathbb{P} \in \Pi$. For every $\mathbb{P} \in \bar{\Pi}$ there is a sequence $\mathbb{P}_n \in \Pi$ converging to $\mathbb{P}$ and thus $\mathbb{P}(K) \geq \limsup_n \mathbb{P}_n(K) \geq 1 - \epsilon$.

Suppose $\mathbb{P}_n$ is an arbitrary sequence of probability measures in $\bar{\Pi}$. Now consider (by Lemma 2.) a compact metric space (Y, ρ) and a homeomorphism $T : \Omega \to Y$ mapping Ω onto $T(\Omega)$. For each B Borel subset in Y, $T^{-1}(B)$ is a Borel subset in Ω. Define

$$\nu_n(B) := \mathbb{P}_n(T^{-1}(B)), \qquad \forall B \in \mathbb{B}(Y)$$

then ν_n is a probability measure on $(Y, \mathbb{B}(Y))$ for each n. Now, by Lemma 1. we see that $\mathcal{P}(Y)$ is compact and hence there is a subsequence ν_{n_k} of ν_n and a probability measure $\nu \in \mathcal{P}(Y)$ such that ν_{n_k} converges weakly to ν as $k \to \infty$. Set $Y_0 := T(\Omega)$, then ν is concentrated on Y_0 (i.e. there exists a set $E \in \mathbb{B}(Y)$ such that $E \subset Y_0$ and $\nu(E) = 1$). Indeed, for each $m \in \mathbb{N}$ take K_m compact sets in Ω such that $\mathbb{P}(K_m) > 1 - \frac{1}{m}$, $\forall \mathbb{P} \in \Pi$. $T(K_m)$ is compact in Y, $\forall m \in \mathbb{N}$, hence

$$\nu(T(K_m)) \geq \limsup_k \nu_{n_k}(T(K_m)) \geq \limsup_k \mathbb{P}_{n_k}(K_m) \geq 1 - \frac{1}{m}$$

Now set $E = \bigcup_{m=1}^{\infty} K_m$ and observe that $\nu(E) \geq \nu(K_m)$ for all m implying $\nu(E) = 1$.
Finally define $\nu_0(A) := \nu(A \cap E)$ for every $A \in \mathbb{B}(Y_0)$. This is a finite Borel measure on Y_0 and $\nu_0(E) = 1$. Define

$$\mathbb{P}(A) := \nu_0(T(A)), \qquad \forall A \in \mathbb{B}(\Omega)$$

This is a Borel probability measure on Ω. Suppose F is a closed set in Ω, then $T(F)$ is closed in $T(\Omega) = Y_0$, hence

$$\limsup_k \mathbb{P}_{n_k}(F) = \limsup_k \nu_{n_k}(T(F)) \leq \nu(T(F)) =$$

$$\nu(T(F) \cap E) + \nu(T(F) \cap E^c) = \nu(T(F) \cap E) = \nu_0(T(F)) = \mathbb{P}(F)$$

this says that $\mathbb{P}_{n_k}$ converges weakly to $\mathbb{P}$ and thus $\bar{\Pi}$ is compact.

Note 1.9. Prohorov's theorem is valid in every seperable and complete metric space and in fact the reverse statement which is the most useful in what follows does not require completeness.

Relatively compactness of a sequence is not sufficient for convergence and what we are going to use further in many cases, is the uniqueness of the subsequental limit point to conclude convergence of sequences. This relies on *Theorem 2.6 from Billingshley [1999]* which is necessary and sufficient condition for weak convergence of a sequence of probability measures that we will use more than once.

Theorem 1.10. A sequence of probability measures $\mathbb{P}_n$ converges weakly to a probability measure $\mathbb{P}$ if and only if every subsequence $\mathbb{P}_{n_k}$ contains further convergent subsequence $\mathbb{P}_{n_{k_m}}$ that converges to $\mathbb{P}$.

Proof: ($\Longrightarrow$) If $\mathbb{P}_n \Rightarrow \mathbb{P}$, then each subsequence converges to the unique limit $\mathbb{P}$.
($\Longleftarrow$) Suppose that $\mathbb{P}_{n_k}$ does not converges to $\mathbb{P}$. Then there exists a bounded and continuous function f such that

$$\int f d\mathbb{P}_n \nrightarrow \int f d\mathbb{P}$$

this implies that there is an $\epsilon > 0$ and a subsequence $\mathbb{P}_{n_k}$ such that

$$\left| \int f d\mathbb{P}_{n_k} - \int f d\mathbb{P} \right| > \epsilon \qquad \text{for all } k$$

but then, no further subsequence can converge to $\mathbb{P}$.

This set up is appropriate only for continuous processes. In order to include stochastic processes with jumps we have to consider a bigger sampling space such as the space of right (or left) continuous functions with left (right) limits. Call D the space of *cadlag* functions

$$D[0,\infty)^d := \{\omega : [0,\infty) \to \mathbb{R}^d / \lim_{s\uparrow t} \omega(s) \quad \text{exists} \quad \text{and} \quad \lim_{s\downarrow t} \omega(s) = \omega(t)\}$$

Every element in D is a function with at most countably many discontinuities. To see this, consider for each $n \in \mathbb{N}$, the set $A_n := \{t > 0 : |\omega(t) - \omega(t^-)| > \frac{1}{n}\}$. Fix n and assume that A_n contains a limit point $\lim_k t_k$. Then there is an increasing subsequence t_{k_m} such that

$$|\omega(t_{k_m}) - \omega(t_{k_m}^-)| > \frac{1}{n}, \quad \forall m$$

this contradicts with the fact that ω has left limits, meaning that for each n, A_n has no limit points, therefore the set of all discontinuities of ω, $\cup_n A_n$ is at most countable.

In order to have similar definition for convergence of probability measures we have to equip D with a metric that makes the space complete and seperable. There is a topology which makes D Polish. This topology is called Skorohod topology and is characterized by the following concept.

A sequence of *cadlag* functions in $[0, \infty)$, ω_n will converge to a point $\omega \in D$ if and only if there exists a sequence of strictly increasing continuous functions λ_n from $[0, \infty)$ onto $[0, \infty)$ with $\lambda(0) = 0$ and $\lim_t \lambda_n(t) = \infty$ such that

$$\sup_t |\lambda_n(t) - t| \to 0$$

and

$$\sup_{t \leq N} |\omega_n(t) - \omega(\lambda_n(t))| \to 0$$

for every $N \geq 1$.

Skorohod topology is weaker than the uniform on compact sets topology. If $\omega_n \subset D$, $n \in \mathbb{N}$, converge to $\omega \in D$ uniformly on compact, then choosing $\lambda_n(t) = t$ we get ω_n converge to ω in Skorohod topology. For continuous functions Skorohod topology coincide with uniform on compacts topology.

Remark 1.11. If ω is a continuous function of $[0, \infty)$, then ω_n converges to ω in Skorohod topology if and only if ω_n converges to ω uniformly on compacts.
Proof: Suppose λ_n be a strictly increasing continuous process such that $\omega_n \to \omega$ in Skorohod. Then

$$|\omega_n(t) - \omega(t)| \leq |\omega_n(t) - \omega(\lambda_n(t))| + |\omega(\lambda_n(t)) - \omega(t)|$$

since ω is continuous, it is uniformly continuous on all compact subsets of $[0, \infty)$ thus,

$$\sup_{t \leq N} |\omega(\lambda_n(t)) - \omega(t)| \underset{n}{\to} 0$$

moreover, by assumption $\sup_{t\leq N}|\omega_n(t)-\omega(\lambda_n(t))| \underset{n}{\to} 0$ and so we get $\omega_n \to \omega$ uniformly on compacts.

In order to define a topology with the above property, first set

$$K_N(t) = \begin{cases} 1, & t \leq N \\ N+1-t, & t \in (N, N+1) \\ 0, & t \geq N+1 \end{cases}$$

and

$$||\lambda||^{\circ} := \sup_{s<t} \left| \ln \frac{\lambda(t)-\lambda(s)}{t-s} \right|$$

For $\omega, \omega' \in D$, define

$$\delta_n(\omega, \omega') = \inf_{\lambda \in \Lambda} \{||\lambda||^{\circ} + ||K_N(\lambda)\omega(\lambda) - K_N\omega'||_{\infty}\}$$

then, the metric δ defined as

$$\delta(\omega, \omega') = \sum_{N=1}^{\infty} \frac{1}{2^N} 1 \wedge \delta_N(\omega, \omega')$$

is called Skorohod metric and it can be shown that makes D Polish *(Theorem 16.3 Billingshley [1999])*.

Compact sets in $D[0,\infty)^d$ are characterized again through Arzelá-Ascoli, but now we can not use Bourbaki's *Theorem X.5.2.* since it is suitable only for continuous functions. The following is *Theorem 16.5.* from *Billingshley [1999]*.

Theorem 1.12. (Arzelá -Ascoli in $D[0,\infty)^d$) A set $A \subset D[0,\infty)^d$ is relatively compact if and only if, for each $n \in \mathbb{N}$

$$\sup_{\omega \in A} \sup_{t \leq n} |\omega(t)| < \infty$$

and

$$\lim_{\delta \to 0} \sup_{\omega \in A} w'_n(\omega, \delta) = 0$$

where

$$w'_n(\omega, \delta) = \inf_{\{t_i\}_\delta} \max_{1 \leq i \leq u} \sup_{s,t \in [t_{i-1}, t_i)} |\omega(s) - \omega(t)| = \inf_{\{t_i\}_\delta} \max_{1 \leq i \leq u} w_\omega[t_{i-1}, t_i)$$

while $\{t_i\}_{i=1}^{u}$ is a partition of $[0, n)$ with the property $t_i - t_{i-1} > \delta$ for every $1 \leq i \leq u - 1$.

Now, If we consider $(\mathcal{P}(D), \delta_P)$, to be the set of probability measures on $(D, \mathbb{B}(D))$ endowed with the Prohorov metric then, Remark 1.4. and Prohorov's Theorem remain valid.

For now on, unless it is stated otherwise, the sampling space will be $C[0, \infty)^d = \Omega$ equipped with the uniform on compacts metric.

We next define the function X that maps every element of Ω to itself, then $X_t(\omega)$ denotes the composition of X with the natural projection $\pi_t : C([0, \infty])^d \to \mathbb{R}^d$ $\pi_t(\omega) = \omega(t)$ meaning

$$X(\omega(t)) = \omega(t)$$

We call the process X the canonical process and the σ-algebra generated by this process, the canonical σ-algebra, this is $\mathcal{G} := \sigma(X_t, t \geq 0)$. A filtration $\{\mathcal{F}_t\}$ on a measurable space is an increasing sequence of sub-σ-algebras, and a process Z is said to be adapted to the filtration if for each time t, the random variable Z_t is measurable with respect to $\mathcal{F}_t$. In our case the smallest filtration under the canonical process X is adapted to; is the natural filtration generated by the process itself i.e. the collection $\{\mathcal{G}_t\}_{t \geq 0} := \sigma(X_s, s \leq t)_{t \geq 0}$. The σ-algebra $\mathcal{G}$ is in fact the same as the Borel σ-algebra $\mathbb{B}(C[0, \infty)^d)$ with respect to the uniform on compacts topology.

Proposition 1.13. $\mathcal{G} = \sigma(X_t, t \geq 0) = \mathbb{B}(C[0, \infty)^d)$

Proof: That $\mathcal{G} \subset \mathbb{B}(C[0, \infty)^d)$ is because π_t is continuous and therefore uniform continuous on compact sets. To prove the other direction we use the seperability of $C[0, \infty)^d$. In particular the collection $\mathcal{C} = \{B(\omega_n, \frac{1}{n}) : k, n \in \mathbb{N}\}$ is a basis for the uniform on compact sets topology, where $\{\omega_n\}$ is a sequence of polynomials that form a countable dense subset of Ω. Meaning, we can write any open subset G with respect to uniform on compact topology as a countable union of elements from the collection $\mathcal{C}$ i.e.

$$G = \bigcup_{n=1}^{\infty} \{\omega \in \Omega : \sup_{t \in [0, T_n]} |\omega(t) - \omega_n(t)| < \frac{1}{n}\}$$

By the continuity of $\omega(t)$ we can take the supremum over the rationals of $[0, T_n]$. And then we can conclude that any open set G is a countable union of $\mathcal{F}$- measurable functions, hence $\mathcal{F}$-measurable.

Definition 1.14. Set $\mathcal{G}_t^+ := \cap_{\epsilon>0}\mathcal{G}_{t+\epsilon}$, if $\mathcal{G}_t^+ = \mathcal{G}_t$ for all t, then the filtration is right-continuous.

Although the process has continuous paths, the natural filtration is not right-continuous. Consider for example, the event

$$A := \{\omega \in \Omega : X(\omega) \text{ is right differentiable at } 0\}$$

then $A \in \mathcal{G}_0^+$ but not in $\mathcal{G}_0 = \{\emptyset, \Omega\}$.

Let $\mathbb{P}$ be a probability measure on $(\Omega, \mathcal{G})$, we complete the σ-algebra $\mathcal{G}_0$ w.r.t. $\mathbb{P}$, i.e. we consider the σ-algebra consisting from sets $A \cup N$, with $A \in \mathcal{G}_0$ and $N \subset M$ while $\mathbb{P}(M) = 0$. A reason why we want to do this, is to gain measurability properties for random variables such as the $\sup_{s\leq t} f(X_s)$ or stopping times $T := \inf\{t \geq 0 : X_t \in A\}$, for some Borel measurable function f and some Borel subset A of $\mathbb{R}^d$.

Suppose $\mathbb{P}$ is a probability measure on $(\Omega, \mathcal{F})$ where $\mathcal{F}$ is the canonical σ-algebra and suppose $\mathcal{F}_t$ is the right-continuous modification of the completion of the natural filtration.

Definition 1.15. A stochastic process X on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is called progressively measurable (progressive) if for each $t > 0$ the random variable X_t is $\mathcal{B}([0,t]) \times \mathcal{F}_t$ measurable. This means

$$\{(s,\omega) \in [0,t] \times \Omega : X(s,\omega) \in A\} \in \mathcal{B}([0,t]) \times \mathcal{F}_t$$

for every Borel subset A in $\mathbb{R}^d$.

The canonical process is $\{\mathcal{F}_t\}_t-$ progressive. This is beacause it is adapted to the filtration $\mathcal{F}_t$ and has continuous paths. It would be enough to have only right or only left continuous paths *(Proposition I.1.13 Karatzas & Shreve [1991])* implying that the canonical process on $D[0,\infty)^d$ is also progressive.

Proposition 1.16. The canonical process X on $D[0,\infty)^d$ is progressively measurable with respect to $\mathcal{F}_t$

Proof: Fix $t > 0$. Define for each $n \in \mathbb{N}$ the process

$$X_s^n(\omega) := X_{\frac{(k+1)t}{2^n}}(\omega) \text{ for } k = 0, 1, ..., 2^n - 1 \text{ and } \frac{kt}{2^n} \leq s \leq \frac{(k+1)t}{2^n}$$

and set $X_0^n(\omega) = X_0(\omega)$. The function $(s,\omega) \mapsto X_s^n(\omega)$ is $\mathbb{B}([0,t]) \otimes \mathcal{F}_t$-measurable, since

$$\{(s,\omega) : X_s^n(\omega) \in A\} = \{(s,\omega) : X_{\frac{(k+1)t}{2^n}}(\omega) \in A\}$$

is either $[0,t]\otimes\{\omega : X_{\frac{(k+1)t}{2^n}}(\omega)\in A\}$ or $\emptyset\otimes\{\omega : X_{\frac{(k+1)t}{2^n}}(\omega)\in A\}$ and in both cases, (because X is adapted to $\mathcal{F}_t$), is an element of $\mathbb{B}([0,t])\otimes\mathcal{F}_t$. Now by the right continuity of the process we have

$$\lim_n X_s^n(\omega) = X_s(\omega) \text{ for } (s,\omega)\in[0,t]\times\Omega$$

therefore X_t is also $\mathbb{B}([0,t])\otimes\mathcal{F}_t$-measurable as a limit of $\mathbb{B}([0,t])\otimes\mathcal{F}_t$-measurable functions.

Definition 1.17. A stochastic process X has the Markov property and is called **Markov process** if

$$\mathbb{E}[X_t|\mathcal{F}_s] = \mathbb{E}[X_t|X_s] \qquad \forall t\geq s$$

Definition 1.18. A stochastic process X_t has the **strong Markov property** if

$$\mathbb{E}[X_{t+T}|\mathcal{F}_T] = \mathbb{E}[X_t|X_T]$$

for every finite stopping time T

Definition 1.19. A stochastic process X_t is a **martingale** with respect to a filtration $\{\mathcal{F}_t\}$ if

- X is integrable
- X is adapted to the filtration $\{\mathcal{F}_t\}_{t\geq 0}$
- $\mathbb{E}[X_t|\mathcal{F}_s] = X_s \qquad \forall t\geq s$

One-dimensional **Brownian Motion** is a martingale. This is a continuous process W_t starting at 0, with independent increments $W_t - W_s$ having normal distribution $N(0,t-s)$ if $s<t$. The d-dimensional Brownian Motion is the d-dimensional martingale process $(W_t^{(1)},...,W_t^{(d)})$ with independent components where each component is one-dimensional Brownian Motion.
The stochastic process X_t is a local martingale if there exists an increasing sequence of stopping times T_n such that $\lim_n T_n=\infty$ and the stopped process $X_{T_n\wedge t}$ is a martingale for each n. If X_t is a local martingale, the quadratic variation of X is defined to be the unique adapted increasing continuous process $\langle X\rangle_t$ such that the process $X_t^2-\langle X\rangle_t$ is a local martingale. In general, square integrable continuous martingales have finite quadratic variation,unbounded first variation and zero higher variations and in case of Borwnian motion the quadratic variation is $\langle W\rangle_t = t$. If X_t and Y_t are $\mathcal{F}_t$-local martingales then the cross variation of X and Y is defined to be the unique adapted continuous process of bounded variation $\langle X,Y\rangle_t$ such that

the process $X_tY_t - \langle X, Y\rangle_t$ is a continuous local martingale.
Suppose M_t is an $\mathcal{F}_t$ - local martingale and X_t is a measurable adapted process satisfying the condition $\mathbb{P}\big(\int_0^t X_s^2 d\langle M\rangle_s < \infty\big) = 1$ for every $t \geq 0$. The stochastic integral of X w.r.t. M is defined to be the unique local martingale process $I_t = \int_0^t X_s dM_s$ such that

$$\langle I, N\rangle_t = \int_0^t X_s d\langle M, N\rangle_s$$

for every N continuous martingale adapted to the filtration $\mathcal{F}_t$.

A continuous semimartingale is a submartingale process that arise as the summation of a continuous local martingale and of an adapted non decreasing process of bounded variation. Every submartingale has a unique a.s. semimartingale decomposition due to Doob and Mayer.
It$\hat{o}$'s formula says that the class of continuous semimartingales is closed under composition with twice differentiable functions and the a.s. unique decomposition of $\{f(X_t)\}_t$ is the following

$$f(X_t) = f(X_0) + \sum_{i=1}^d \int_0^t \partial_i f(X_s) dX_s^{(i)} + \frac{1}{2}\sum_{i,j=1}^d \int_0^t \partial_{ij} f(X_s) d\langle X^{(i)}, X^{(j)}\rangle_s$$

Brownian Motion plays important role in stochastic analysis and has the following characterization due to *P.Lev$\grave{y}$ [1948]*. (*Theorem III.3.16. Karatzas & Shreve [1991]*)

Theorem 1.20. *(Lev$\grave{y}$'s characterization of Brownian Motion)* If M is a continuous d-dimensional $\mathcal{F}_t$ local martingale and $\langle M^{(i)}, M^{(j)}\rangle_t = \delta_{ij} t$, for every $1 \leq i, j \leq d$, then the process M is a d-dimensional Brownian Motion.

Proof: By definition, we have to prove that the increments $M_t - M_s$ are independent of $\mathcal{F}_s$ and have d-dimensional normal distribution with 0 mean and covariance matrix $(t-s)\mathbb{I}_d$, for each $0 \leq s < t$. In order to do this, it is sufficient to show for every $u \in \mathbb{R}^d$ $\mathbb{P}$-a.s.

$$\mathbb{E}\Big[e^{i<u,x>}\Big|\mathcal{F}_s\Big] = \mathbb{E}\Big[e^{i<u,x>}\Big] = e^{-\frac{1}{2}||u||^2(t-s)}$$

Fix $u \in \mathbb{R}^d$ and consider $f(x) = e^{i<u,x>}$. Then $\partial_j f(x) = iu_j f(x)$ and $\partial_{jk} f(x) = -u_j u_k f(x)$. We apply It$\hat{o}$'s formula to f.

$$e^{i<u,X_t>} = e^{i<u,X_s>} + \sum_{j=1}^d \int_s^t iu_j e^{i<u,X_r>} dM_r^{(j)} - \frac{1}{2}\sum_{j=1}^d \int_s^t {u_j}^2 e^{i<u,X_r>} dr$$

Because M is continuous square integrable martingale and f is bounded by 1, $\int_0^t f(X_r)dM_r^{(j)}$ is a continuous square integrable martingale and thus,

$$\mathbb{E}\left[\int_s^t e^{i<u,X_r>}dM_r^{(j)}\middle|\mathcal{F}_s\right] = 0 \qquad \mathbb{P}-\text{a.s.}$$

Given $A \in \mathcal{F}_s$, we multiply $e^{i<u,X_t>}$ by $e^{-i<u,X_s>}\mathbb{1}_A$ we have

$$\begin{aligned} e^{i<u,X_t-X_s>}\mathbb{1}_A = \mathbb{1}_A + \mathbb{1}_A \sum_{j=1}^{d} \int_s^t iu_j e^{i<u,X_r-X_s>}dM_r^{(j)} \\ -\frac{1}{2}\mathbb{1}_A||u||^2 \int_s^t e^{i<u,X_r-X_s>}dr \end{aligned}$$

we now take expectations and use Fubini's theorem to exchange integrals,

$$\mathbb{E}[e^{i<X_t-X_s>}\mathbb{1}_A] = \mathbb{P}(A) - \frac{1}{2}||u||^2 \int_s^t \mathbb{E}[e^{i<u,X_r-X_s>}\mathbb{1}_A]dr$$

This is a deterministic equation (Volterra integral equation of second order) explicitly solved

$$\mathbb{E}[e^{i<u,X_t-X_s>}\mathbb{1}_A] = \mathbb{P}(A)e^{-\frac{1}{2}||u||^2(t-s)}$$

which is what we wanted since u in $\mathbb{R}^d$ and A in $\mathcal{F}_s$ was chosen arbitrary.

Stochastic integrals of measurable adapted processes w.r.t. Brownian Motion are local martingales if $\mathbb{P}(\int_0^t X_s ds < \infty) = 1$. The converse is the Martingale representation theorem

Theorem 1.21. Suppose M is a martingale related to a filtration on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. There is an extended probability space $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}})$ which a d-dimensional Brownian Motion W independent of M and a progressively measurable $(d \times d)$ matrix process ξ live in such that

$$\mathbb{E}\left[\int_0^t \xi_s^2 ds < \infty\right]$$

for every $0 \leq t < \infty$
and the martingale process M has the following representation

$$M^{(i)} = \sum_{j=1}^{d} \int_0^t \xi_s^{(ij)} dW_s^j$$

The last result in the introductory chapter is Girsanov's transformation *[1960]*. The theorem states that in a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ equipped with a filtration which satisfies the usual conditions, the class of continuous martingale processes lie in $(\Omega, \mathcal{F}, \mathbb{P})$, remain invariant under a change of measure *Cameron & Martin formula [1944]* given that the changed probability measure is absolutely continuous with $\mathbb{P}$.

Let M be a positive continuous martingale starting from 1. Fix $T > 0$, we define on $(\Omega, \mathcal{F})$ a new probability measure Q_T by setting, for each $0 \leq t \leq T$, M_t to be the Radon - Nikodym derivative $\frac{dQ_T}{d\mathbb{P}}$ on the events of $\mathcal{F}_t$. i.e.

$$Q_T(A) = \mathbb{E}_{\mathbb{P}}[M_t \mathbb{1}_A]$$

whenever $A \in \mathcal{F}_t$. We now, define a probability measure Q on $(\Omega, \mathcal{F})$ that restricted to any $\mathcal{F}_t$, agrees with Q_t. Such a probability measure does not always exists. Since the family of measures Q_t is consitence, as a result of the martingale property, the existence of such a probability measure Q on $(\mathbb{R}^{d^{[0,\infty)]}}, \mathbb{B}(\mathbb{R}^{d^{[0,\infty)]}}))$ is due to Kolmogorov consistency theorem (*Theorem II.2.2. Karatzas & Shreve [1991]*), but we will need this result and Girsanov theorem for the space of continuous functions and the following existence/extension theorem is the desired result in spaces of right-continuous functions (*Theorem, Families of Consistent Probability Measures. A.S. Cherny [2002]*).

Proposition 1.22. Suppose Q_t is a consistent family of probability measures on $(\Omega, \mathcal{F}_t)$. Then there exists a probability measure Q on $(\Omega, \mathcal{F})$ such that $Q|_{\mathcal{F}_t} = Q_t$.

Proof: We start by proving the statement for the natural canonical filtration $\mathcal{G}_t = \sigma(X_s, s \leq t)$, then we can replace $\mathcal{G}_t$ with it's right continuous modification using the following argument. Suppose we find a probability measure Q that restricted to $\mathcal{G}_t$ coincides with Q_t, then

$$Q|_{\mathcal{F}_t} = Q|_{\mathcal{G}_{t+1}}|_{\mathcal{F}_t} = Q_{t+1}|\mathcal{F}_t = Q_t$$

Let Q_n be a consistent sequence of probability measures defined on $(\Omega, \mathcal{G}_n)$. We define for each $n \in \mathbb{N}$, the probability measures μ_n such that $\mu_n(A) = Q_n(X_{\cdot \wedge n} \in A)$. Now each μ_n is a measure on $(\Omega, \mathcal{G})$ where $\mathcal{G}$ is the Borel $\mathbb{B}(C[0,\infty)^d)$ (by Proposition 1.13.). By Remark 1.7. every singleton probability measure on $(\Omega, \mathcal{G})$ is tight. Using Theorem 3.6. this is equivalent to the following two conditions.
1. for each $\epsilon > 0$, there is a $\zeta > 0$ and a $n_0 \in \mathbb{N}$ such that $\mu_n(|X_0| > \zeta) < \epsilon$,

for every $n \geq n_0$.
2. for each $\epsilon > 0$ and $\eta > 0$, there is a $\delta > 0$ and a positive integer n_0 such that

$$\mu_n(w_x(\delta) > \eta) < \epsilon$$

for every $n \geq n_0$. Where $w_x(\delta)$ is the modulus of continuity of the coordinate process.
So each μ_n satisfies the above two conditions. Fix $N \in \mathbb{N}$, then there are $\delta > 0, \zeta > 0 and \eta > 0$ such that

$$\mu_N(|X_0| > \zeta) < \epsilon$$

and

$$\mu_N\Big(\sup_{\substack{|t-s|<\delta \\ t,s \leq N}} |X_t - X_s| > \eta\Big) < \epsilon$$

Now we use the consistency. Because Q_n agree with Q_N on the events of $\mathcal{G}_N$ and because N is arbitrary, we conclude that the two conditions in Theorem 3.6. hold for $n_0 = N$. This yields that the sequence μ_n is tight. By Prohorov's Theorem, μ_n is relatively compact and thus it contains a convergent subsequence μ_{n_k}. The subsequental limit point μ, will be the desired probability measure.

Fix $M > 0$, and consider the map $G : C[0, \infty)^d \to C[0, M]^d$, with $G(\omega) = \omega(t)$, $t \leq M$. Take $\tilde{k}$ to be the minimal k such that $n_k \geq M$ and define probability measures R and R_k, for $k \geq \tilde{k}$

$$R = \mu \circ G^{-1} \qquad \text{and} \qquad R_k = \mu_{n_k} \circ G^{-1}$$

If $k \geq \tilde{k}$, $R_k = R_{\tilde{k}}$. Because G is continuous, R_{n_k} converges weakly to R and thus $R_{\tilde{k}} = R$. So, on the event of $G^{-1}(\mathbb{B}(C[0, M]^d))$ the measures $\mu_{n_{\tilde{k}}}$ an Q coincide, but $G^{-1}(\mathbb{B}(C[0, M]^d)) = \mathcal{G}_M$, hence

$$Q|_{\mathcal{G}_M} = \mu_{n_{\tilde{k}}}|_{\mathcal{G}_M} = \mu_M|_{\mathcal{G}_M} = Q_M$$

Theorem 1.23. (Girsanov) Suppose X, M are continuous square integrable martingales and $M_0 = 0$. Suppose also that the stochastic exponential

$$N_t := \exp(M_t - \frac{1}{2}\langle M \rangle_t)$$

is a true martingale under $\mathbb{P}$ and let Q be the unique probability measure such that the restriction $\frac{dQ}{d\mathbb{P}} := N_t$. Then the stochastic process

$$X_t - \langle X, M \rangle$$

is a martingale process under Q and has the same quadratic variation with X under $\mathbb{P}$.

Proof: Assume without loss of generality that $X_0 = 0$ and let $A \in \mathcal{F}_s$. Then

$$\mathbb{E}_Q[X_t \mathbb{1}_A] = \mathbb{E}_{\mathbb{P}}[N_t X_t \mathbb{1}_A]$$

using integration by parts formula for stochastic integrals this equals to

$$\mathbb{E}_{\mathbb{P}}\Big[\mathbb{1}_A \int_0^t N_r dX_r\Big] + \mathbb{E}_{\mathbb{P}}\Big[\mathbb{1}_A \int_0^t X_r dN_r\Big] + \mathbb{E}_{\mathbb{P}}[\langle X, N\rangle_t \mathbb{1}_A]$$

the above two stochastic integrals are well defined proper martingales therefore

$$\mathbb{E}_Q[X_t \mathbb{1}_A] = \mathbb{E}_{\mathbb{P}}\Big[\mathbb{1}_A \int_0^s N_r dX_r\Big] + \mathbb{E}_{\mathbb{P}}\Big[\mathbb{1}_A \int_0^s X_r dN_r\Big] + \mathbb{E}_{\mathbb{P}}[\langle X, N\rangle_t \mathbb{1}_A] =$$

$$\mathbb{E}_Q[X_s \mathbb{1}_A] - \mathbb{E}_{\mathbb{P}}[\langle X, N\rangle_s \mathbb{1}_A] + \mathbb{E}_{\mathbb{P}}[\langle X, N\rangle_t \mathbb{1}_A]$$

next we have

$$\mathbb{E}_Q[\{\langle X, N\rangle_t - \langle X, N\rangle_s\}\mathbb{1}_A] = \mathbb{E}_{\mathbb{P}}[N_t\{\langle X, N\rangle_t - \langle X, N\rangle_s\}\mathbb{1}_A] =$$

$$\mathbb{E}_{\mathbb{P}}\Big[\mathbb{1}_A \int_s^t \mathcal{E}(M)_r d\langle X, M\rangle_r\Big] = \mathbb{E}_{\mathbb{P}}\Big[\mathbb{1}_A \int_s^t d\langle X, M\rangle_r\Big] =$$

$$\mathbb{E}_{\mathbb{P}}[\{\langle X, N\rangle_t - \langle X, N\rangle_s\}\mathbb{1}_A]$$

hence the process $X - \langle X, M\rangle$ is a martingale with quadratic variation same under both measures.

1.2 Weak and Strong Solutions

Suppose $(\Omega, \mathcal{F}, \mathbb{P})$ is a probability space and $b : \mathbb{R}^d \to \mathbb{R}^d$, $\sigma : \mathbb{R}^d \to \mathbb{R}^d \times \mathbb{R}^d$ are Borel measurable functions and W is a d-dimensional Brownian motion. A stochastic process X who satisfies the integral equation

$$X_t - X_0 = \int_0^t b(X_s)ds + \int_0^t \sigma(X_s)dW_s, \qquad \forall t \geq 0$$

$$X_0 = x \qquad \mathbb{P}\text{-a.s.}$$

is a solution to the stochastic differential equation for b, σ with initial condition $x \in \mathbb{R}^d$. The above integral equation can be written in a differential form without a differential meaning

$$dX_t = b(X_t)dt + \sigma(X_t)dW_t$$

This solution is also called pathwise solution to the SDE and the functions b and σ are called drift and diffusion coefficients respectively.
A solution to SDE(b, σ, x) is called pathwise unique if there exist a stochastic process X that satisfies $dX_t = b(X_t)dt + \sigma(X_t)dW_t$ and if Y is another solution to SDE(b, σ, x) then $X = Y$ $\mathbb{P}-$a.s. . There are other two definitions of solutions to the stochastic differential equation

Definition 1.24. A pathwise solution X of the SDE for b, σ, x, W is called a **strong solution** if X is adapted to the filtration generated by the Brownian motion W

Definition 1.25. The pair (X, W) is called **weak solution** to the SDE for b, σ, x if X, W are both live on $(\Omega, \mathcal{F}, \mathbb{P})$, $X_0 = x$ $\mathbb{P}$-a.s. and X satisfies the integral equation

$$dX_t = b(X_t)dt + \sigma(X_t)dW_t \qquad \mathbb{P}\text{-a.s.}$$

A weak solution (X, W) is said to be weakly unique, if (X', W') is another weak solution, then, under $\mathbb{P}$, X' has the same probability law with X and W' is a Brownian Motion.

Note 1.26. If the SDE has a strong solution, then the SDE has a weak solution.
The converse is not always true. One counterexample is due to Tanaka (*Example 5.3.5 Karatzas & Shreve [1991]*). We consider the function sgn to be

$$sgn(x) = \begin{cases} -1, & x \leq 0 \\ 1, & x > 0 \end{cases}$$

then we get, for each $x \in \mathbb{R}$, $sgn(x) = \frac{1}{sgn(x)}$.

Example 1.27. The one-dimensional equation $dX_t = sgn(X_t)dW_t$ $(*)$ admits a weak but not a strong solution.
Proof: Suppose B_t is one-dimensional Brownian Motion on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. We define the process

$$W_t := \int_0^t sgn(B_s)dB_s$$

This is a stochastic integral of a progressively measurable process with respect to Brownian Motion with $\mathbb{E}[\int_0^t sgn^2(B_s)ds] = t < \infty$, therefore W_t is a continuous square integrable martingale with quadratic variation same as

Brownian Motion. By Lévy's characterization theorem 1.5., W_t is also one-dimensional Brownian Motion. This and the fact that $dB_t = sgn(B_t)dW_t$ imply that (B, W) is a weak solution to $(*)$.
If (B', W') is another weak solution, then, using the same argument, B' is again Brownian Motion, implying that the weak solution to $(*)$ is weakly unique. We see that if (X, W) is a weak solution to $(*)$,then we have also

$$d(-X_t) = sgn(-X_t)dW_t$$

So if the solution is pathwise unique then $X = -X$ $\mathbb{P}-$a.s. This contradicts with the fact that X is Brwonian Motion and therefore spends zero time at zero, meaning that pathwise uniqueness does not hold.
Finally suppose that a strong solution X exists for equation $(*)$ i.e. X is a pathwise solution to $dX_t = sgn(X_t)dW_t$ that is adapted to $\mathcal{F}_t^W$; which is the filtration generated by the Brownian Motion W. Call $\mathcal{F}_t^X$ the filtration generated by the process X itself. Then, since $\mathcal{F}_t^X$ is the smallest filtration X is adapted to, we get $\mathcal{F}_t^X \subset \mathcal{F}_t^W$ for every $0 \leq t < \infty$. Next using the Tanaka formula (*Proposition 3.6.8. Karatzas & Shreve [1991]*)

$$W_t = \int_0^t sgn(X_s)dX_s = |X_t| - \lim_{\epsilon \to 0} \frac{1}{2\epsilon}\lambda(\{s \leq t : |X_s| \leq \epsilon\})$$

where λ is the Lebesque measure. By this we get that W is adapted to $\mathcal{F}_t^X$, meaning $\mathcal{F}_t^X \subset \mathcal{F}_t^W \subset \mathcal{F}_t^{|X|}$ for each $0 \leq t < \infty$ which does not hold and thus $(*)$ does not admit a strong solution.

Suppose X is a pathwise solution to the SDE for b, σ, x, W. *Itô's* formula says that whenever f is an element of $C^2(\mathbb{R}^d)$ then

$$f(X_t) - f(X_0) = \int_0^t \frac{1}{2}\sum_{i,j=1}^d a_{ij}(X_s)\partial_{ij}f(X_s) + \sum_{i=1}^d b_i(X_s)\partial_i f(X_s)ds$$

$$+ \int_0^t \sum_{i=1}^d \sigma_i(X_s)\partial_i f(X_s)dW_s$$

where a is the symmetric matrix function $\sigma\sigma^T$.
The integrant quantity of the Riemann-Stieltjes integral is a linear second order partial defferential operator $\mathcal{L}$ acting on X_s. Thus the process

$$f(X_t) - f(X_0) - \int_0^t \mathcal{L}f(X_s)ds$$

is a local martingale with respect to $\{\mathcal{F}_t\}_{t\geq 0}$ for every $f \in C^2(\mathbb{R}^d)$.

1.3 Martingale Problem

We saw that if X is a solution to the stochastic differential equation for b, σ, and x then by *Itô* formula the process

$$f(X_t) - f(X_0) - \int_0^t \mathcal{L}f(X_s)ds$$

is a local martingale for every $f \in C^2(\mathbb{R}^d)$. If in addition σ is bounded then the above process is a true martingale.
The idea of Stroock & Varadhan is to use this as a key element for characterization of diffusions

Definition 1.28. Let E be seperable and let $\mathcal{A} : \mathbf{B}(E) \to \mathbf{B}(E)$ be a linear operator. A stochastic basis $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$ equipped with an $\{\mathcal{F}_t\}$-measurable process X is a solution to the martingale problem for $\mathcal{A}$ with initial condition $\mu \in \mathcal{P}(E)$ if

$$f(X_t) - f(X_0) - \int_0^t \mathcal{A}f(X_s)ds$$

is a martingale with respect to the filtration $\{\mathcal{F}_t\}$, for every f in the domain of $\mathcal{A}$, $\mathbb{P}$-a.s. and $\mathbb{P} \circ X_0^{-1} = \mu$

In what follows we will use a definition suited to continuous processes that also associates martingales and SDE's. We doing that, first by taking $(\Omega, \mathcal{F})$ to be the canonical measurable space while $\mathcal{F}$ and $\{\mathcal{F}_t\}$ satisfy the usual conditions and second by replacing the arbitrary linear operator with the infinitesimal generator of a Markov semigroup and consider martingale problems that arise from this operator.

Suppose $b : \mathbb{R}^d \to \mathbb{R}^d$ and $a : \mathbb{R}^d \to \mathbb{R}^d \times \mathbb{R}^d$ are measurable functions and let $\mathcal{L}$ be the differential operator acting on functions of $C^2(\mathbb{R}^d)$

$$\mathcal{L}f(x) = \frac{1}{2}\sum_{i,j=1}^{d} a_{ij}\partial_{ij}f(x) + \sum_{i=1}^{d} b_i(x)\partial_i f(x)$$

Definition 1.29. A probability measure $\mathbb{P}$ is a **solution to the martingale problem for the operator $\mathcal{L}$ with initial condition** x if

$$f(X_t) - f(X_0) - \int_0^t \mathcal{L}f(X_s)ds$$

is a martingale for every $f \in C^2(\mathbb{R}^d)$ and

$$\mathbb{P}(X_0 = x) = 1$$

Definition 1.30. Uniqueness for the matingale problem for the operator $\mathcal{L}$ with initial condition x is said to hold if whenever $\mathbb{P}$ and $\mathbb{P}'$ are two solutions to MP$(\mathcal{L}, \delta_x)$ then $\mathbb{P} = \mathbb{P}'$.

The martingale problem is **well-posed** if there exists a unique solution for every $x \in \mathbb{R}^d$.

Being restricted on $C[0,\infty)^d$ if the martingale problem for $\mathcal{L}$ is well posed for every initial condition $x \in \mathbb{R}^d$ then the martingale problem for $\mathcal{L}$ with initial condition some probability measure μ of $\mathbb{R}^d$ is well posed.

Proposition 1.31. If the martingale problem for the operator $\mathcal{L}$ is well posed for every $x \in \mathbb{R}^d$ then the martingale problem for $\mathcal{L}$ with initial condition $\mu \in \mathcal{P}(\mathbb{R}^d)$ is well posed.

Proof: We fix $x \in \mathbb{R}^d$. Suppose that $\mathbb{P}^x$ is the unique solution of MP$(\mathcal{L}, \delta_x)$. We define on Ω a new probability measure $\mathbb{P}^\mu$ mapping each cylindrical Borel set A to $\int_{\mathbb{R}^d} \mathbb{P}^x(A)\mu(dx)$. Then

$$\mathbb{P}^\mu(X_0 \in A) = \mu(A)$$

we name $M_t^f := f(X_t) - f(X_0) - \int_0^t \mathcal{L}f(Xs)ds$ and we take $B \in \mathcal{F}_s$, then

$$\mathbb{E}_{\mathbb{P}^\mu}[M_t^f \mathbb{1}_B] = \int_{\mathbb{R}^d} \mathbb{E}_{\mathbb{P}^x}[M_t^f \mathbb{1}_B]\mu(dx) = \int_{\mathbb{R}^d} \mathbb{E}_{\mathbb{P}^x}[M_s^f \mathbb{1}_B]\mu(dx) = \mathbb{E}_{\mathbb{P}'}[M_s^f \mathbb{1}_B]$$

Suppose now that there are two probability measure $\mathbb{P}_1, \mathbb{P}_2$ as solutions to the martingale problem for $\mathcal{L}$ with initial condition μ. For $\Gamma \in \mathcal{F}$ and $s \geq 0$ we define

$$\tilde{\mathbb{P}}_1(\Gamma) := \mathbb{E}_{\mathbb{P}_1}[\mathbb{E}_{\mathbb{P}^x}[\mathbb{1}_\Gamma|\mathcal{F}_s]]$$

and

$$\tilde{\mathbb{P}}_2(\Gamma) := \mathbb{E}_{\mathbb{P}_2}[\mathbb{E}_{\mathbb{P}^x}[\mathbb{1}_\Gamma|\mathcal{F}_s]]$$

For the new probability measures we have

$$\tilde{\mathbb{P}}_1(X_0 = x) = \mathbb{E}_{\mathbb{P}_1}[\mathbb{E}_{\mathbb{P}^x}[\mathbb{1}_{\{X_0 = x\}}|\mathcal{F}_s]] = 1$$

and

$$\tilde{\mathbb{P}}_2(X_0 = x) = \mathbb{E}_{\mathbb{P}_2}[\mathbb{E}_{\mathbb{P}^x}[\mathbb{1}_{\{X_0 = x\}}|\mathcal{F}_s]] = 1$$

For every $B \in \mathcal{F}_u$, $t \geq u$. If $s < u < t$

$$\mathbb{E}_{\tilde{\mathbb{P}}_1}[M_t^f \mathbb{1}_B] = \mathbb{E}_{\mathbb{P}_1}[\mathbb{E}_{\mathbb{P}^x}[M_t^f \mathbb{1}_B|\mathcal{F}_s]] = \mathbb{E}_{\mathbb{P}_1}[\mathbb{E}_{\mathbb{P}^x}[\mathbb{E}_{\mathbb{P}^x}[M_t^f \mathbb{1}_B|\mathcal{F}_s]|\mathcal{F}_u]]$$

$$= \mathbb{E}_{\mathbb{P}_1}[\mathbb{E}_{\mathbb{P}^x}[\mathbb{E}_{\mathbb{P}^x}[M_t^f \mathbb{1}_B|\mathcal{F}_u]|\mathcal{F}_s]] = \mathbb{E}_{\mathbb{P}_1}[\mathbb{E}_{\mathbb{P}^x}[M_u^f \mathbb{1}_B|\mathcal{F}_s]] = \mathbb{E}_{\tilde{\mathbb{P}}_1}[M_u^f \mathbb{1}_B]$$

If $u < s < t$

$$\mathbb{E}_{\tilde{\mathbb{P}}_1}[M_t^f \mathbb{1}_B] = \mathbb{E}_{\mathbb{P}_1}[\mathbb{E}_{\mathbb{P}^x}[M_t^f \mathbb{1}_B|\mathcal{F}_s]] = \mathbb{E}_{\mathbb{P}_1}[\mathbb{1}_B \mathbb{E}_{\mathbb{P}^x}[M_t^f|\mathcal{F}_s]]$$

$$= \mathbb{E}_{\mathbb{P}_1}[\mathbb{1}_B M_s^f] = \mathbb{E}_{\mathbb{P}_1}[M_u^f \mathbb{1}_B] = \mathbb{E}_{\tilde{\mathbb{P}}_1}[M_u^f \mathbb{1}_B]$$

Finally if $u < t < s$

$$\mathbb{E}_{\tilde{\mathbb{P}}_1}[M_t^f \mathbb{1}_B] = \mathbb{E}_{\mathbb{P}_1}[\mathbb{E}_{\mathbb{P}^x}[M_t^f \mathbb{1}_B|\mathcal{F}_s]] = \mathbb{E}_{\mathbb{P}_1}[M_t^f \mathbb{1}_B] = \mathbb{E}_{\mathbb{P}_1}[M_u^f \mathbb{1}_B] = \mathbb{E}_{\tilde{\mathbb{P}}_1}[M_u^f \mathbb{1}_B]$$

meaning that the process M_t^f is an $\mathcal{F}_t$-martingale under measure $\tilde{\mathbb{P}}_1$. In the same way we see that the measure $\tilde{\mathbb{P}}_2$ is also a solution to the martingale problem $\mathrm{MP}(\mathcal{L}, \delta_x)$. By the uniqueness of the $\mathrm{MP}(\mathcal{L}, \delta_x)$ we get $\tilde{\mathbb{P}}_1 = \tilde{\mathbb{P}}_2$. Because s was arbitrary we get

$$\mathbb{E}_{\mathbb{P}_1}[\mathbb{E}_{\mathbb{P}^x}[\mathbb{1}_\Gamma|\mathcal{F}_s]] = \mathbb{E}_{\mathbb{P}_2}[\mathbb{E}_{\mathbb{P}^x}[\mathbb{1}_\Gamma|\mathcal{F}_s]]$$

for every $\Gamma \in \mathcal{F}$ and for every $s \geq 0$ which implies $\mathbb{P}_1(\Gamma) = \mathbb{P}_2(\Gamma)$ for every $\Gamma \in \mathcal{F}$.

Next we see that uniqueness of the martingale problem for the differential operator $\mathcal{L}$ implies the Markov property *(Theorem 4.4.2. Ethier & Kurtz[1986])*. Later we will see that also implies the Strong Markov property.

Proposition 1.32. Suppose that for every two solutions of the martingale problem for $\mathcal{L}$ with initial condition $x \in \mathbb{R}^d$, we have that the cannonical process X has the same one-dimensional distribution both under the two probability measures (i.e.)

$$\mathbb{P}[X_t \in A] = \mathbb{P}'[X_t \in A] \qquad \forall A \in \mathcal{F}, \forall t \geq 0$$

then X has the same family of finite-dimensional distributions under both measures meaning that $\mathrm{MP}(\mathcal{L}, \delta_x)$ is unique. Additionally X is a Markov process under $\mathbb{P}^x$, for every $x \in \mathbb{R}^d$.

Proof: Let $\mathbb{P}, \mathbb{P}'$ be solutions of martingale problem as described. In order to show that the cannonical process X has the same family of finite dimensional distributions under those measures it is sufficient to show that

$$\mathbb{E}_{\mathbb{P}}\left[\prod_{i=1}^{n} g_i(X_{t_i})\right] = \mathbb{E}_{\mathbb{P}'}\left[\prod_{i=1}^{n} g_i(X_{t_i})\right] \tag{0.3}$$

for every bounded continuous functions g_i, for every $t_i \in [0, \infty)$, for every $n \in \mathbb{N}$. By the uniqueness hypothesis for $n = 1$ (1) holds. Proceeding by strong induction, we assume that (1) holds for every integer $1 \leq k \leq n$. Fixing $0 \leq t_1 < ... < t_n$ we define two new probability measures on the same measurable space

$$\mathbf{Q}(A) = \frac{\mathbb{E}_{\mathbb{P}}[\mathbb{1}_A \mathbb{1}_{\{A \notin \mathcal{F}_{t_n}\}} \prod_{i=1}^{n} g_i(X_{t_i})]}{\mathbb{E}_{\mathbb{P}}[\prod_{i=1}^{n} g_i(X_{t_i})]} + \mathbb{E}_{\mathbb{P}}[\mathbb{1}_A \mathbb{1}_{A \in \mathcal{F}_{t_n}}]$$

and

$$\mathbf{Q}'(A) = \frac{\mathbb{E}_{\mathbb{P}'}[\mathbb{1}_A \mathbb{1}_{\{A \notin \mathcal{F}_{t_n}\}} \prod_{i=1}^{n} g_i(X_{t_i})]}{\mathbb{E}_{\mathbb{P}'}[\prod_{i=1}^{n} g_i(X_{t_i})]} + \mathbb{E}_{\mathbb{P}'}[\mathbb{1}_A \mathbb{1}_{A \in \mathcal{F}_{t_n}}]$$

The new probability measures $\mathbf{Q}, \mathbf{Q}'$ are solutions of martingale problem for operator $\mathcal{L}$ with initial condition $x \in \mathbb{R}^d$.

$$\mathbf{Q}(X_0 = x) = \mathbb{P}(X_0 = x) = 1$$

and

$$\mathbf{Q}'(X_0 = x) = \mathbb{P}'(X_0 = x) = 1$$

Suppose $B \in \mathcal{F}_s$

$$\mathbb{E}_{\mathbf{Q}}[M_t^f \mathbb{1}_B] = \frac{\mathbb{E}_{\mathbb{P}}[M_t^f \mathbb{1}_B \mathbb{1}_{\{M_t^f \mathbb{1}_B \notin \mathcal{F}_{t_n}\}} \prod_{i=1}^{n} g_i(X_{t_i})]}{\mathbb{E}_{\mathbb{P}}[\prod_{i=1}^{n} g_i(X_{t_i})]} + \mathbb{E}_{\mathbb{P}}[M_t^f \mathbb{1}_B \mathbb{1}_{\{M_t^f \mathbb{1}_B \in \mathcal{F}_{t_n}\}}]$$

In case $s < t_n < t$ we get

$$\mathbb{E}_{\mathbf{Q}}[M_t^f \mathbb{1}_B] = \mathbb{E}_{\mathbb{P}}[M_t^f \mathbb{1}_B] = \mathbb{E}_{\mathbb{P}}[M_s^f \mathbb{1}_B] = \mathbb{E}_{\mathbf{Q}}[M_s^f \mathbb{1}_B]$$

If $t_n < s < t$

$$\begin{aligned} &\mathbb{E}_{\mathbf{Q}}[M_t^f \mathbb{1}_B] \\ &= \frac{\mathbb{E}_{\mathbb{P}}[M_t^f \mathbb{1}_B \mathbb{1}_{\{M_t^f \mathbb{1}_B \notin \mathcal{F}_{t_n}\}} \prod_{i=1}^{n} g_i(X_{t_i})]}{\mathbb{E}_{\mathbb{P}}[\prod_{i=1}^{n} g_i(X_{t_i})]} \\ &= \frac{\mathbb{E}_{\mathbb{P}}[\mathbb{E}_{\mathbb{P}}[M_t^f \mathbb{1}_B \prod_{i=1}^{n} g_i(X_{t_i}) | \mathcal{F}_s]]}{\mathbb{E}_{\mathbb{P}}[\prod_{i=1}^{n} g_i(X_{t_i})]} \end{aligned}$$

$$= \frac{\mathbb{E}_{\mathbb{P}}[\mathbb{E}_{\mathbb{P}}[M_s^f \mathbb{1}_B \prod_{i=1}^n g_i(X_{t_i})]}{\mathbb{E}_{\mathbb{P}}[\prod_{i=1}^n g_i(X_{t_i})]}$$
$$= \mathbb{E}_{\mathbf{Q}}[M_s^f \mathbb{1}_B]$$

Finally if $s < t < t_n$

$$\mathbb{E}_{\mathbf{Q}}[M_t^f \mathbb{1}_B] = \mathbb{E}_{\mathbb{P}}[M_t^f \mathbb{1}_B] = \mathbb{E}_{\mathbb{P}}[M_s^f \mathbb{1}_B] = \mathbb{E}_{\mathbf{Q}}[M_s^f \mathbb{1}_B]$$

Similarly calculation for $\mathbf{Q}'$ shows that $\mathbf{Q}, \mathbf{Q}' \in \mathrm{MP}(\mathcal{L}, \delta_x)$. Furthermore for $k = n$ we get

$$\mathbb{E}_{\mathbf{Q}}[h(X_t)] = \mathbb{E}_{\mathbf{Q}'}[h(X_t)]$$

for every bounded continuous function h, $t \geq 0$. We choose $t_{n+1} > t_n$, the above relation implies

$$\mathbb{E}_{\mathbb{P}}[h(X_{t_{n+1}}) \prod_{i=1}^n g_i(X_{t_i})] = \mathbb{E}_{\mathbb{P}'}[h(X_{t_{n+1}}) \prod_{i=1}^n g_i(X_{t_i})]$$

This means that (1) holds for $k = n + 1$ and the solution to the martingale problem $\mathrm{MP}(\mathcal{L}, \delta_x)$ is a unique probability measure $\mathbb{P}^x$.
For the Markov property we need to show

$$\mathbb{E}_{\mathbb{P}^x}[f(X_{t+s})|\mathcal{F}_s] = \mathbb{E}_{\mathbb{P}^x}[f(X_{t+s})|\sigma(X_s)]$$

for every f bounded and continuous function.
We fix $r \geq 0$ and choose an element of $\mathcal{F}_r$, F with $\mathbb{P}^x(F) > 0$. We define probability measures as follow

$$\mathbb{P}_1(A) = \frac{\mathbb{E}_{\mathbb{P}^x}[\mathbb{1}_F \mathbb{E}_{\mathbb{P}^x}[\mathbb{1}_A|\mathcal{F}_r]]}{\mathbb{P}^x(F)}$$

and

$$\mathbb{P}_2(A) = \frac{\mathbb{E}_{\mathbb{P}^x}[\mathbb{1}_F \mathbb{E}_{\mathbb{P}^x}[\mathbb{1}_A|X_r]]}{\mathbb{P}^x(F)}$$

As before, $X_0 = x$ under both $\mathbb{P}_1$ and $\mathbb{P}_2$ and for $B \in \mathcal{F}_s$

$$\mathbb{E}_{\mathbb{P}_1}[M_t^f \mathbb{1}_B] = \frac{\mathbb{E}_{\mathbb{P}^x}[\mathbb{1}_F \mathbb{E}_{\mathbb{P}^x}[M_t^f \mathbb{1}_B|\mathcal{F}_r]]}{\mathbb{P}^x(F)}$$

If $s < r$, because $B \cap F \in \mathcal{F}_s$, we get

$$\mathbb{E}_{\mathbb{P}_1}[M_t^f \mathbb{1}_B] = \frac{\mathbb{E}_{\mathbb{P}^x}[\mathbb{1}_F M_r^f \mathbb{1}_B]}{\mathbb{P}^x(F)} = \frac{\mathbb{E}_{\mathbb{P}^x}[\mathbb{1}_F \mathbb{E}_{\mathbb{P}^x}[M_s^f \mathbb{1}_B|\mathcal{F}_r]]}{\mathbb{P}^x(F)} = \mathbb{E}_{\mathbb{P}_1}[M_s^f \mathbb{1}_B]$$

and if $s > r$, using tower property of conditional expectation we get

$$\mathbb{E}_{\mathbb{P}_1}[M_t^f \mathbb{1}_B] = \frac{\mathbb{E}_{\mathbb{P}^x}[\mathbb{1}_F \mathbb{E}_{\mathbb{P}^x}[\mathbb{E}_{\mathbb{P}^x}[M_t^f \mathbb{1}_B|\mathcal{F}_s]|\mathcal{F}_r]]}{\mathbb{P}^x(F)} = \mathbb{E}_{\mathbb{P}_1}[M_s^f \mathbb{1}_B]$$

Similarly for $\mathbb{P}_2$. Meaning again that the measures $\mathbb{P}_1$,$\mathbb{P}_2$ are solutions to MP$(\mathcal{L}, \delta_x)$. By the uniqueness of the martingale problem, we conclude that $\mathbb{P}_1 = \mathbb{P}_2 = \mathbb{P}^x$. This implies that

$$\mathbb{E}_{\mathbb{P}^x}[\mathbb{1}_F \mathbb{E}_{\mathbb{P}^x}[f(X_t)|\mathcal{F}_r]] = \mathbb{E}_{\mathbb{P}^x}[\mathbb{1}_F \mathbb{E}_{\mathbb{P}^x}[f(X_t)|X_r]]$$

for every non-null $\mathcal{F}_r$-measurable set F, for every bounded and continuous f. Because $r \geq 0$ was fixed arbitrary we have the Markov property.

Now that we have the main definition, we will study a compactness criterion involving martingale problems. The next theorem *(Theorem 1.4.6. Stroock & Varadhan [1979])* gives necessary and sufficient conditions for tightness of a family of probability measures on space of continuous functions. The relation with tightness comes from Arzelà - Ascoli and again Prohorov's Theorem translates the tightness of the family to compactness of it's closure.

Fix $\eta > 0$. For $n \in \mathbb{N}$, define exit times

$$\tau_n = \inf\{t \geq \tau_{n-1} : |X_t - X_{\tau_{n-1}}| \geq \frac{\eta}{4}\}$$

and

$$\tau_0 = 0$$

Fix $T > 0$. Define $N := \inf\{n \in \mathbb{N} : \tau_{n+1} > T\}$ and $\delta(\eta) := \min_{n \leq N}\{\tau_n - \tau_{n-1}\}$

Under the following conditions *Stroock & Varadhan* obtained an estimate for the modulus of continuity of the canonical process.

(1). For all non-negative functions $f \in C_0^\infty(\mathbb{R}^d)$, there is a constant $A_f \geq 0$ such that the process $f(X_t) + A_f t$ is a non-negative submartingale.
(2). Given a function $f \in C_0^\infty(\mathbb{R}^d)$, the choice of A_f can be made so it works for all translates of f

Lemma 1.33. Let $(E, \mathcal{G}, \mathbb{P})$ be a probability space and $\{\xi_n\}_{n \in \mathbb{N}}$ be a non-decreasing sequence of random variables taking values on $[0, \infty]$. Suppose that ξ is adapted to a filtration $\mathcal{G}_n$ and additionally suppose there exists a $\lambda < 1$, such that

$$\mathbb{E}_{\mathbb{P}}\left[e^{-(\xi_{n+1}-\xi_n)}\middle|\mathcal{G}_n\right] \leq \lambda$$

then if $T > 0$ and $N = \inf\{n \geq 0 : \xi_{n+1} > T\}$, then $N < \infty$ and

$$\mathbb{P}(N > k) \leq e^T \lambda^k$$

Proof:

$$\mathbb{P}(N > k) = \mathbb{P}(\xi_k \leq T) = \mathbb{E}_{\mathbb{P}}[\mathbb{1}_{\{\xi_k \leq T\}} e^{-\xi_k} e^{\xi_k} \leq e^T \mathbb{E}_{\mathbb{P}}[e^{-\xi_k}]$$

also by the assumptions

$$\mathbb{E}_{\mathbb{P}}[e^{-\xi_{k+1}} | \mathcal{G}_k] = e^{-xi_k} \mathbb{E}_{\mathbb{P}}\left[e^{-(\xi_{k+1}-\xi_k)} \middle| \mathcal{G}_k\right] \leq e^{-\xi_k}\lambda$$

and thus, by induction

$$\mathbb{E}_{\mathbb{P}}[e^{-\xi_k}] \leq \lambda^k$$

Lemma 1.34. Under conditions (1) and (2) and for any $n \geq 0$, on the event $\{\tau_n < \infty\}$ $\mathbb{P}-$a.s.

$$\mathbb{P}(\tau_{n+1} - \tau_n \leq \delta | \mathcal{F}_{\tau_n}) \leq A_{\frac{\eta}{4}} \delta$$

Proof: Set $\epsilon := \frac{\eta}{4}$. Consider a $C_0^\infty(\mathbb{R}^d)$ function $0 \leq f \leq 1$ such that $f(0) = 1$ and $f(x) = 0$, if $|x| \geq \epsilon$. By hypothesis, the process $G(t) := f(X_t) + A_\epsilon t$ is a submartingale under $\mathbb{P}$ w.r.t. $\mathcal{F}_t$. Now, given $s \leq t$, take $A \in \mathcal{F}_s$ and $B \in \mathcal{F}_{\tau_n}$. Then

$$\mathbb{E}_{\mathbb{P}}[\mathbb{E}_{\mathbb{P}}[G(t)\mathbb{1}_A | \mathcal{F}_{\tau_n}]\mathbb{1}_{B\cap\{\tau_n \leq s\}}] = \mathbb{E}_{\mathbb{P}}[G(t)\mathbb{1}_{A\cap B\cap\{\tau_n \leq s\}}] \geq$$

$$\mathbb{E}_{\mathbb{P}}[G(s)\mathbb{1}_{A\cap B\cap\{\tau_n \leq s\}}] = \mathbb{E}_{\mathbb{P}}[\mathbb{E}_{\mathbb{P}}[G(s)\mathbb{1}_A | \mathcal{F}_{\tau_n}]\mathbb{1}_{B\cap\{\tau_n \leq s\}}]$$

this implies

$$\mathbb{E}_{\mathbb{P}}[G(t)\mathbb{1}_A | \mathcal{F}_{\tau_n}] \geq \mathbb{E}_{\mathbb{P}}[G(s)\mathbb{1}_A | \mathcal{F}_{\tau_n}] \quad \mathbb{P}-\text{a.s. on } \{\tau_n \leq s\}$$

Take $\omega' \in F^c \cap \{\tau_n \leq s\}$, where $\mathbb{P}(F) = 0$ such that

$$\mathbb{E}_{Q^{\omega'}}[G(t)\mathbb{1}_A] := \mathbb{E}_{\mathbb{P}}[G(t)\mathbb{1}_A | \mathcal{F}_{\tau_n}](\omega') \geq \mathbb{E}_{\mathbb{P}}[G(s)\mathbb{1}_A | \mathcal{F}_{\tau_n}](\omega') := \mathbb{E}_{Q^{\omega'}}[G(s)\mathbb{1}_A]$$

then

$$\mathbb{E}_{Q^{\omega'}}[G(t)\mathbb{1}_A\mathbb{1}_{\{\tau_n \leq t\}}] \geq \mathbb{E}_{Q^{\omega'}}[G(t)\mathbb{1}_A\mathbb{1}_{\{\tau_n \leq s\}}] \geq \mathbb{E}_{Q^{\omega'}}[G(s)\mathbb{1}_A\mathbb{1}_{\{\tau_n \leq s\}}]$$

meaning $\{f(X_t) + A_\epsilon t\}\mathbb{1}_{\{\tau_n \leq t\}}$ is a submartingale process under $Q^{\omega'}$ for each $\omega' \in F^c$.

Next, define $f^{\omega'}(X_t) := f(X_t - X_{\tau_n(\omega')}(\omega'))$ on $\{\tau_n(\omega') < \infty\}$ and $f(X_t) = 1$, otherwise. We also have the submartingale property for the stopped process and the stopping time τ_{n+1}, on $\{\tau_n(\omega') < \infty\}$. In particular,

$$\mathbb{E}_{Q^{\omega'}}[f^{\omega'}(X_{\tau_{n+1}\wedge\tau_n(\omega')+\delta}) + A_\epsilon(\tau_n(\omega') + \delta)] \geq \mathbb{E}_{Q^{\omega'}}[f^{\omega'}(X_{\tau_n(\omega')}) + A_\epsilon \tau_n(\omega')]$$

because $f(0) = 1$, this is

$$\mathbb{E}_{Q^{\omega'}}[f^{\omega'}(X_{\tau_{n+1}\wedge\tau_n(\omega')+\delta}) + A_\epsilon \delta] \geq 1$$

for any $\omega' \in F^c$. This is

$$\mathbb{E}_{Q^{\omega'}}[1 - f^{\omega'}(X_{\tau_{n+1}\wedge\tau_n(\omega')+\delta})] \leq A_\epsilon \delta$$

Finally, on the event $\{\tau_{n+1} - \tau_n(\omega') \leq \delta\}$ inside $\{\tau_n(\omega') < \infty\}$, by the definitions of the exit times and f

$$1 - f^{\omega'}(X_{\tau_{n+1}\wedge\tau_n(\omega')+\delta}) = 1 - f(X_{\tau_{n+1}} - X_{\tau_n(\omega')}(\omega'))1 - 0 = 1$$

This says, that on $\{\tau_n(\omega') < \infty\}$ for any ω' in F^c while $\mathbb{P}(F) = 0$

$$\mathbb{E}_{Q^{\omega'}}[\mathbb{1}_{\{\tau_{n+1}-\tau_n(\omega')\leq\delta\}}] \leq A_\epsilon \delta$$

which is the same with

$$\mathbb{E}_{\mathbb{P}}[\mathbb{1}_{\tau_{n+1}} - \tau_n \leq \delta | \mathcal{F}_{\tau_n}] \leq A_\epsilon \delta \quad \mathbb{P} - \text{a.s. on}\{\tau_n < \infty\}$$

Theorem 1.35. Let Π be a family of probability measures on $(\Omega, \mathcal{F})$ such that

$$\lim_{\zeta\to\infty} \sup_{\mathbb{P}\in\Pi} \mathbb{P}(|X_0| \geq \zeta) = 0$$

Additionally, assume that every probability measure $\mathbb{P}$ in Π, satisfies the conditions (1) and (2) and that the choice of the constant A_f can be made independent of $\mathbb{P}$. Then Π is relatively compact.

Proof: By Theorem 7.3. Billingshley, it is sufficient to show, for all $\eta > 0$ for all $T > 0$

$$\lim_{\delta\to 0} \sup_{\mathbb{P}\in\Pi} \mathbb{P}\left(\sup_{\substack{|t-s|\leq\delta \\ t,s\leq T}} |X_t - X_s| > \eta \right) = 0$$

First, note that $\{\delta(\eta) > \delta\} \subset \left\{ \sup_{|t-s|<\delta} |X_t - X_s| \leq \eta \right\}$. To see this, consider an ω such that $\delta_\omega(\eta) > \delta$ and take $t, s \leq T$ such that $|t-s| < \delta < \delta_\omega(\eta)$. The sub-intervals $[\tau_{n-1}(\omega), \tau_n(\omega))$ have length greater than $\delta_\omega(\eta)$ except possibly

$[\tau_N, T]$, since $\delta_\omega(\eta)$ defined to be the minimum such length. Therefore t, s are in the same sub-interval or in adjacent sub-intervals. Therefore, in case t, s lie in the same sub-interval $[\tau_{n-1}, \tau_n)$

$$|X_t(\omega) - X_s(\omega)| \leq |X_t(\omega) - X_{\tau_{n-1}(\omega)}(\omega)| + |X_s(\omega) - X_{\tau_{n-1}(\omega)}(\omega)| \leq \frac{\eta}{2}$$

and in any case,

$$|X_t - X_s| \leq \eta$$

implying the inclusion

$$\sup_{|t-s|<\delta<\delta_\omega(\eta)} |X_t(\omega) - X_s(\omega)| \leq \eta$$

Instead, it is sufficient to show

$$\lim_{\delta \to 0} \sup_{\mathbb{P} \in \Pi} \mathbb{P}\Big(\delta(\eta) \leq \delta\Big) = 0$$

We use the conditions (1) and (2) through Lemma 1.34.

$$\begin{aligned} \mathbb{P}(\delta(\eta) \leq \delta) \leq \mathbb{P}(\min_{n \leq k}\{\tau_n - \tau_{n-1}\}) + \mathbb{P}(N > k) \leq \\ \sum_{n=1}^{k} \mathbb{P}(\tau_n - \tau_{n-1} \leq \delta) + \mathbb{P}(N > k) \leq \\ k\delta A_{\frac{\eta}{4}} + \mathbb{P}(N > k) \end{aligned}$$

thus, we need to show

$$\lim_{k \to \infty} \sup_{\mathbb{P} \in \Pi} \mathbb{P}(N > k) = 0$$

This holds again by Lemma 1.34.

$$\mathbb{E}_{\mathbb{P}}\Big[e^{-(\tau_{n+1}-\tau_n)}\Big|\mathcal{F}_{\tau_n}\Big] \leq \mathbb{P}(\tau_{n+1} - \tau_n \leq t_0|\mathcal{F}_{\tau_n}) + e^{-t_0}\mathbb{P}(\tau_{n+1} - \tau_n > t_0|\mathcal{F}_{\tau_n})$$

$$= e^{-t_0} + (1 - e^{-t_0})\mathbb{P}(\tau_{n+1} - \tau_n \leq t_0|\mathcal{F}_{\tau_n}) \leq e^{-t_0} + (1 - e^{-t_0})A_{\frac{\eta}{4}}t_0$$

Call $\lambda := e^{-t_0} + (1 - e^{-t_0})A_{\frac{\eta}{4}}t_0$ and choose t_0 such that $\lambda < 1$, then by Lemma 1.33.

$$\sup_{\mathbb{P} \in \Pi} \mathbb{P}(N > k) \leq e^T \lambda^k \underset{k \to \infty}{\to} 0$$

2 Equivalence between Well Posedness for Martingale Problem and Well Posedness for SDE

Suppose the stochastic differential equation with coefficients b and σ has a weak solution which is unique in the sense of probability law, then we say that the SDE is well posed. This unique solution induces on $(C[0,\infty)^d, \mathcal{B}(C[0,\infty)^d]))$ a measure which solves the local martingale problem for the associated operator. The converse is also true and in fact on $C[0,\infty)^d$ there is a complete equivalence between weak solution for SDE and solutions for the assiciated (local) martingale problem*(Theorem 4.5.2. Stroock & Varadhan[1979]).*
The coefficients a_{ij}, b_i, σ are supposed to be measurable functions. First we consider the case where the operator $\mathcal{L}$ is uniformly elliptic *(Theorem V.1.1. R.Bass[1997])* i.e.

$$\sup_i ||b_i||_\infty < B$$

$$A|y|^2 < \sum_{i,j=1}^{d} a_{ij}(x) y_i y_j < A^{-1}|y|^2$$

for A,B some positive real numbers and for every $x, y \in \mathbb{R}^d$

Theorem 2.1. Let $a = \sigma\sigma^T$ and $\mathcal{L}$ is uniformly elliptic. The stochastic differential equation for b,σ and starting point x has a weak solution which is weakly unique if and only if the associated (local)martingale problem has a unique solution.
Proof: **Existence**
($\Longrightarrow$) Suppose the SDE for b, σ and x has a weak solution. This means that

$$X_t = X_0 + \int_0^t b(X_s)ds + \int_0^t \sigma(X_s)dW_s$$

for some d-dimensional Brownian motion W. We apply Itô formula

$$f(X_t) = f(X_0) + \sum_{i=1}^{d} \int_0^t \partial_i f(X_s) dX_s^{(i)} + \frac{1}{2}\sum_{i,j=1}^{d} \int_0^t \partial_{ij} f(X_s) d\langle X^{(i)}, X^{(j)}\rangle_s$$

Calculating the variation of every coordinate of X and the covariation between the coordinates we get

$$dX_t^{(i)} = b_i(X_t)dt + \sum_{j=1}^{d} \sigma_{ij}(X_t) dW_t^{(j)}$$

$$d\langle X^{(i)}, X^{(j)}\rangle_s = \sum_{n=1}^{d}\sum_{m=1}^{d} \sigma_{in}(X_t)\sigma_{jm}(X_t) d\langle W^{(n)} W^{(m)}\rangle_t$$

The d-dimensional Brownian motion is a vector of d independent one-dimensional Brownian motions meaning $d\langle W^{(n)} W^{(m)}\rangle_t = 0$ whenever $n \neq m$. So

$$d\langle X^{(i)}, X^{(j)}\rangle_s = \sum_{n=1}^{d} \sigma_{in}(X_t)\sigma_{jn}(X_t) dt = a_{ij}(X_t)dt$$

So we get

$$f(X_t) = f(X_0) + \int_0^t b(X_s)\partial_i f(X_s) + \frac{1}{2} a_{ij}(X_s)\partial_{ij} f(X_s) ds$$

$$+ \int_0^t \sum_{i,j=1}^{d} \partial_i f(X_s)\sigma_{ij}(X_s) dW_s^{(j)}$$

$$f(X_t) = f(X_0) + \int_0^t \mathcal{L} f(X_s) ds + \int_0^t \sum_{i,j=1}^{d} \partial_i f(X_s)\sigma_{ij}(X_s) dW_s^{(j)}$$

If the functions b, σ are supposed to be bounded and f has compact support the stochastic integral becomes true martingale. So the distribution $\mathbb{P}$ of the process X is a solution to the martingale problem for $\mathcal{L}$ with initial condition x.

($\Longleftarrow$) Suppose now that $\mathbb{P}$ is a solution to (local)MP($\mathcal{L}, \delta_x$). The test process M_t^f is a local martingale for every $f \in C^2(\mathbb{R}^d)$. We choose $f(x) = x_i$ for every $x \in \mathbb{R}^d$. Then the i-th component of the process M_t^{id} is continuous a (local) martingale

$$M_t^{(i)} := X_t^{(i)} - X_0^{(i)} - \int_0^t b_i(X_s) ds$$

Next for $f(x) = x_i x_j$

$$M_t^{(ij)} := X_t^{(i)} X_t^{(j)} - X_0^{(i)} X_0^{(j)} - \int_0^t \left\{ a_{ij} + b_i(X_s) X_s^{(j)} + b_j(X_s) X_s^{(i)} \right\} ds$$

is also by assumption a continuous local martingale. Therefore we can write the process $M_t^{(i)} M_t^{(j)} - \int_0^t a_{ij}(X_s) ds$ in the form

$$M_t^{(ij)} - M_t^{(i)} X_0^{(j)} - M_t^{(j)} X_0^{(i)} + \int_0^t (X_s^{(i)} - X_t^{(i)}) b_j(X_s) ds + \int_0^t (X_s^{(i)} - X_t^{(i)}) b_j(X_s) ds$$

$$+\int_0^t b_j(X_s)ds\int_0^t b_i(X_s)ds$$

which equals to

$$M_t^{(ij)}-M_t^{(i)}X_0^{(j)}-M_t^{(j)}X_0^{(i)}+\int_0^t(M_s^{(i)}-M_t^{(i)})b_j(X_s)ds+\int_0^t(M_s^{(i)}-M_t^{(i)})b_j(X_s)ds$$

$$-\int_0^t b_j(X_s)\int_s^t b_i(X_r)drds-\int_0^t b_i(X_s)\int_s^t b_j(X_r)drds+\int_0^t b_j(X_s)ds\int_0^t b_i(X_s)ds$$

because the process $M_t^{(i)}$ is a martingale we have that $\int_0^t |b_i(X_s)|ds < \infty$ $\mathbb{P}-$a.s. By Fubini we get that

$$M_t^{(i)}M_t^{(j)} - \int_0^t a_{ij}(X_s)ds = M_t^{(ij)} - M_t^{(i)}X_0^{(j)} - M_t^{(j)}X_0^{(i)}$$

$$+\int_0^t(M_s^{(i)} - M_t^{(i)})b_j(X_s)ds + \int_0^t(M_s^{(i)} - M_t^{(i)})b_j(X_s)ds$$

We can write the term $\int_0^t(M_s^{(i)} - M_t^{(i)})b_j(X_s)ds$ as follows

$$\int_0^t(M_s^{(i)} - M_t^{(i)})b_j(X_s)ds = -\int_0^t\int_s^t dM_u^{(i)}b_j(X_s)ds$$

We then use stochastic Fubini and write

$$-\int_0^t\int_0^u b_j(X_s)dsdM_u^{(i)} = -\int_0^t\int_0^s b_j(X_u)dudM_s^{(i)}$$

This means that the integrals $\int_0^t(M_s^{(i)}-M_t^{(i)})b_j(X_s)ds$ and $\int_0^t(M_s^{(i)}-M_t^{(i)})b_j(X_s)ds$ are both martingales. Because these are Riemann Stieltjes integrals this implies that are both zero. Finally we get that

$$M_t^{(i)}M_t^{(j)} - \int_0^t a_{ij}(X_s)ds = M_t^{(ij)} - M_t^{(i)}X_0^{(j)} - M_t^{(j)}X_0^{(i)}$$

is a sum of martingales, therefore a martingale.
Because the operator is uniformly elliptic the $d \times d$ matrix a is symmetric strictly positive definite and so it's square root is an invertible matrix, thus we can define the process

$$W_t := \int_0^t \sigma^{-1}(X_s)dM_s^{id}$$

The process σ^{-1} is also M^{id} - integrable because

$$\begin{aligned}\mathbb{E}\Big[W_t^{(i)^2}\Big] &= \mathbb{E}\Big[\int_0^t \sum_{n=1}^d \sigma_{in}^{-1}(X_s) \sum_{m=1}^d \sigma_{im}^{-1}(X_s) d\langle M^{(n)}, M^{(m)}\rangle_s\Big] \\ &= \mathbb{E}\Big[\int_0^t \sum_{n,m=1}^d a_{nm}^{-1}(X_s)a_{nm}(X_s)ds\Big] = t < \infty\end{aligned}$$

and

$$\begin{aligned}\mathbb{E}\Big[W_t^{(i)}W_t^{(j)}\Big] &= \mathbb{E}\Big[\int_0^t \sum_{n=1}^d \sigma_{in}^{-1}(X_s) \sum_{m=1}^d \sigma_{jm}^{-1}(X_s) d\langle M^{(n)}, M^{(m)}\rangle_s\Big] \\ &= \mathbb{E}\Big[\int_0^t \sum_{n,m=1}^d \sigma_{in}^{-1}(X_s)\sigma_{jm}^{-1}a_{nm}(X_s)ds\Big] \\ &= \mathbb{E}\Big[\int_0^t \sum_{n,m=1}^d \sigma_{in}^{-1}(X_s)\sigma_{jm}^{-1}\sum_{k=1}^d \sigma_{nk}(X_s)\sigma_{km}(X_s)ds\Big] \\ &= \mathbb{E}\Big[\int_0^t \sum_{m,k=1}^d \sigma_{jm}^{-1}\sigma_{km}(X_s)id_{ik}ds\Big] \\ &= \mathbb{E}\Big[\int_0^t \sum_{m=1}^d \sigma_{jm}^{-1}(X_s)\sigma_{im}(X_s)ds\Big] = \mathbb{E}\Big[\int_0^t id_{ij}ds\Big] = id_{ij}t\end{aligned}$$

That means, the process W is a continuous martingale with $d\langle W^{(i)}, W^{(j)}\rangle_t = \delta_{ij}t$ under the probability measure $\mathbb{P}$.
By Levý characterisation theorem for Brownian motion, W is a d-dimensional Brownian motion. We now recal that the process $M_t^{id} = X_t - X_0 - \int_0^t b(X_s)ds$. Then we have

$$\begin{aligned}dW_t &= \sigma^{-1}(X_t)dM_t^{id} \\ dW_t &= \sigma^{-1}(X_t)\big(dX_t - b(X_t)dt\big) \\ dX_t &= b(X_t)dt + \sigma(X_t)dW_t\end{aligned}$$

Namely, the pair (X, W) is a solution to SDE for b, σ and x under $\mathbb{P}$.
Uniqueness ($\Longrightarrow$) Suppose the solution to the stochastic differential equation is weakly unique. Let $\mathbb{P}_1$ and $\mathbb{P}_2$ are two solutions to the associated martingale problem MP$(\mathcal{L}, \delta_x)$. By the first assertion of the theorem, there exist a d-dimensional Brownian motion W_1 such that

$$dX_t = b(X_t)dt + \sigma(X_t)dW_{1t} \qquad \mathbb{P}_1\text{-a.s.}$$

and similarly, a d-dimensional Brownian motion such that

$$dX_t = b(X_t)dt + \sigma(X_t)dW_{2t} \qquad \mathbb{P}_2\text{-a.s.}$$

By the weak uniqueness assumption, this implies that $\mathbb{P}_1 X^{-1} = \mathbb{P}_2 X^{-1}$ and we have uniqueness for the martingale problem.
($\Longleftarrow$) Conversely, suppose now that the martingale problem MP($\mathcal{L}, \delta_x$) has a unique solution and let (X^1, W_1), (X^2, W_2) be two weak solutions of the associated SDE under a probability measure $\mathbb{P}$ on a filtered probability space $(E, \mathcal{G})$. The d-dimensional process X^1 is continuous and maps E onto Ω. We define the probability measure $\mathbb{P}_1$ on $(\Omega, \mathcal{F})$ to be the measure

$$\mathbb{P}_1(Z \in A) = \mathbb{P}(X^1 \in A)$$

for every A cylindrical set. Similarly we define

$$\mathbb{P}_2(Z \in A) = \mathbb{P}(X^2 \in A)$$

The cannonical process will be Z such that $Z(X^1) = X^1$. Then, $\mathbb{P}_1(Z_0 = x) = \mathbb{P}({X^1}_0 = x) = 1$ and $\mathbb{P}_2(Z_0 = x) = \mathbb{P}({X^2}_0 = x) = 1$. Also the process M^f is a local martingale under both measures $\mathbb{P}_1$ and $\mathbb{P}_2$. By the uniqueness of martingale problem $\mathbb{P}_1 = \mathbb{P}_2$. This means that

$$\mathbb{P}(X^1 \in A) = \mathbb{P}(X^2 \in A)$$

for every A cylindrical set, meaning that every two solutions to SDE have the same finite dimensional distributions i.e. the solution to SDE is weakly unique.

This equivalence is still valid if we consider operators not necessarily uniformly elliptic. In the above proof we use uniform ellipticity in order to define the driving Brownian motion for the stochastic differential equation $dM_t^{id} = \sigma(X_t)dW_t$. We now see that this can be done without the assumption that σ^{-1} exists *(Proposition 4.6 [p.316-317], Karatzas & Shreve[1991]).*

Theorem 2.2. Suppose $a = \sigma\sigma^T$. Let $\mathbb{P}$ is a solution to the martingale problem for $\mathcal{L}$ with initial condition x, then there exists a d-dimensional Brownian motion W on an extension probability space $(E, \mathcal{G}, \mathbb{Q})$ of $(\Omega, \mathcal{F}, \mathbb{P})$ such that the process (X, W) is a weak solution to SDE for b,σ and x.

Proof: The process $M_t^{id} = (M_t^{(1)}, ..., M_t^{(d)})$ where $M_t^{(i)} = X_t^{(i)} - X_0^{(i)} - \int_0^t b_i(X_s)ds$ is a continuous d-dimensional $\mathcal{F}_t$- martingale. The cross covariance between components $\langle M^{(i)}, M^{(j)} \rangle_t$ is the absolutely continuous function over t for almost all ω, $\int_0^t a_{ij}(X_s)ds$. We use the martingale representation theorem to establish the existence of a d-dimensional Brownian motion B on an extention probability space $(E, \mathcal{G}, \mathbb{Q})$ adapted to a filtration $\{\mathcal{G}_t\}_t$ and a $d \times d$ matrix of measurable adapted processes $\{\xi^{ij}{}_t\}_t$ such that

$$\mathbb{E}_{\mathbb{Q}}\Big[\int_0^t \xi^{ij}{}_s^2 ds\Big] < \infty \qquad \text{for all } t \in [0, \infty), \text{ for all } i, j \in \{i, ..., d\}$$

and

$$M_t^{id} = \int_0^t \xi_s dB_s \qquad \mathbb{Q}\text{-a.s.}$$

Next we need to construct an d-dimensional Brownian motion W such that

$$\int_0^t \sigma(X_s)dW_s = \int_0^t \xi_s dB_s \qquad \mathbb{Q}\text{-a.s.}$$

We consider the $d \times d$ matrix function $R(\xi, \sigma)$ defined on the set

$$D = \{(\xi, \sigma) : \xi, \sigma \in M_{d \times d}(\mathbb{R}) \quad \text{and} \quad \xi\xi^T = \sigma\sigma^T\}$$

such that $\sigma R = \xi$ and $RR^T = I$ *(Problem 5.4.7. Karatzas & Shreve [1991]).* Such function R exists because the matrices $\sigma\sigma^T$ and $\xi\xi^T$ are symmetric therefore can be diagonalized with decompositions $U\Lambda_\sigma U^T$ and $V\Lambda_\xi V^T$ where U, V are orthogonal . Consequently we can write $\sigma = \sqrt{\sigma\sigma^T}U$ and $\xi = \sqrt{\xi\xi^T}V$. Next we define $R(\xi, \sigma)$ to be $U^{-1}V$ and check that R satisfy the above conditions

$$\sigma U^{-1}V = \sqrt{\sigma\sigma^T}V = \sqrt{\xi\xi^T}V = \xi$$
$$U^{-1}V(U^{-1}V)^T = U^{-1}VV^TU = I$$

Finally we define the d-dimensional Brownian motion to be the process

$$W_t = \int_0^t R^T(\xi_s, \sigma(X_s))dB_s$$

The process $R^T(\xi_s, \sigma(X_s))$ is $\mathcal{G}_t$-progressive as composition of the Borel measurable R^T and the $\mathcal{G}_t$-progressive $(\xi, \sigma(X))$. This is implies that W a continuous $\{\mathcal{G}_t\}-$martingale w.r.t. $\mathbb{Q}$ since

$$\mathbb{E}_{\mathbb{Q}}\Big[W_t^{(i)^2}\Big] = \mathbb{E}_{\mathbb{Q}}\Big[\int_0^t \sum_{j,k=1}^d R^T{}_{ij}R_{ik}ds\Big] = t < \infty$$

especially

$$\langle W^{(i)}, W^{(j)}\rangle_t = \delta_{ij}t$$

By Levỳ's characterization, W is a Brownian motion and therefore

$$\int_0^t \sigma(X_s)dW_s = \int_0^t \xi_s dB_s$$

Remark 2.3. The $d \times d$ function $R(\xi, \sigma)$ is Borel measurable.

Proof: We consider the characteristic polynomial of $a = \sigma\sigma^T$

$$\chi_a(\lambda) = \lambda^d + c_{d-1}\lambda^{d-1} + ... + c_0$$

The coefficients of χ_a are linear combinations of elements a_{ij}, $0 \leq i, j \leq d$, meaning c_k, $0 \leq k \leq d-1$ are Borel measurable functions. The roots of the polynomial are continuous functions of the coefficients, therefore the eigenvalues of a are Borel measurable functions.
We next use *Theorem 1, Edward A. Azoff [1974]* : If E is a σ-compact, closed set in $X \times Y$, where X, Y are complete seperable metric spaces, then $\pi_X(E)$ is Borel and there exists a Borel measurable function $\phi : \pi_X(E) \to Y$ whose graph is contained in E. Like *Corollary 5, Edward A. Azoff* now define

$$E = \{(a, U, \Lambda) \in M_{d\times d}(\mathbb{R}) \times M_{d\times d}(\mathbb{R}) \times M_{d\times d}(\mathbb{R}) : U^{-1}aU = \Lambda,\ \Lambda \text{ is diagonal}$$

$$\text{U is orthonormal}\}$$

The projection $\pi_{M_{d\times d}(\mathbb{R})\times M_{d\times d}(\mathbb{R})} = M_{d\times d}(\mathbb{R})$, since every symmetric positive semi definite matrix admits diagonalization. Consider Borel measurable function $\phi : M_{d\times d}(\mathbb{R}) \times M_{d\times d}(\mathbb{R}) \to M_{d\times d}(\mathbb{R})$ of Theorem 1 of Azoof and set

$$U(x) = \begin{cases} \phi(a(x), \Lambda(x)), & \text{when is is defined} \\ 0, & \text{elsewhere} \end{cases}$$

Using the same process, the matrix $\xi\xi^T$ will have a Borel diagonalization $V\Lambda_\xi V^T$ and therefore the matrix function $R(\xi, \sigma)$ defined above as $U^{-1}V$ is Borel measurable.

3 Uniqueness of Martingale Problem Solution and Convergence of Probability measures

In this section at first we present a sufficient condition in order to have uniqueness of MP, relating existence of another problem; the Cauchy problem (*Stroock & Varadhan [1969]*). Then, if we assume uniqueness we see that we get as a result a Strong Markov diffusion process. After that, we study convergence of solutions of martingale problems. The basic result is that if the MP for continuous bounded operator $\mathcal{L}$ is well posed with unique solution a probability measure $\mathbb{P}$ and we can appropriately approximate this operator with a sequence $\mathcal{L}_n$ for which the MP admits solution $\mathbb{P}^n$, $\forall n$, then we have convergence of measures $\mathbb{P}^n$ to $\mathbb{P}$ under the weak topology. Later we will use this fact to prove two existence results: The first is existence of solutions of MP for continuous coefficients. The second which is a stronger result, is existence of solutions of MP for measurable coefficients but stays valid only in the elliptic case. In order to prove the last result we will first use the Girsanov theorem to eliminate the neccesity of the drift coefficient.

3.1 Conditions for Uniqueness

Suppose that f is a continuous function with compact support. A bounded continuous function u of $[0, \infty) \times \mathbb{R}^d$ which is continuously differentiable in t and C^2 with bounded first and second derivatives in x, is a solution to Cauchy problem if

$$\begin{cases} \frac{\partial u(t,x)}{\partial t} = \mathcal{L}u(t,x) & , t > 0, x \in \mathbb{R}^d \\ u(0,x) = f(x) & , x \in \mathbb{R}^d \end{cases}$$

Theorem 3.1. *(Stroock & Varadhan [1969])* Suppose that for any $f \in C_0^\infty(\mathbb{R}^d)$ the Cauchy problem has a solution $u_f \in C([0,\infty)\times\mathbb{R}^d) \cap C^{1,2}((0,\infty)\times\mathbb{R}^d)$ which is bounded on each strip $[0,T] \times \mathbb{R}^d$, then if the martingale problem for $\mathcal{L}$ with initial condition $x \in \mathbb{R}^d$ admits a solution, it is unique.
Proof: Let $\mathbb{P}_1^x$ and $\mathbb{P}_2^x$ are two solutions of the martingale problem MP$(\mathcal{L}, \delta_x)$ and $f \in C_0^\infty(\mathbb{R}^d)$. For fixed T we define the function $g(t,x) = u_f(T-t,x)$. g is an element of $C([0,T] \times \mathbb{R}^d) \cap C^{1,2}((0,T) \times \mathbb{R}^d)$ and satispfies

$$\begin{cases} \frac{\partial g(t,x)}{\partial t} + \mathcal{L}g(t,x) = 0 & , t \in (0,T), x \in \mathbb{R}^d \\ g(T,x) = f(x) & , x \in \mathbb{R}^d \end{cases}$$

For each $\mathbb{P}_i^x$, $i = 1, 2$, we consider the probability measure $\mathbb{Q}_i$ on the corresponding extended probability space under which the cannonical process X is a solution to SDE for b, σ and x. We use Itô's formula for $g(t, x)$.

$$
\begin{aligned}
g(t, X_t) - g(0, X_0) = & \int_0^t \frac{\partial g(s, X_s)}{\partial s} + \mathcal{L}g(s, X_s)ds \\
& + \int_0^t \sum_{i,j=1}^{d} \frac{\partial g(t, X_s)}{\partial x_i} \sigma_{ij}(X_s) dW_s^j
\end{aligned}
$$

meaning

$$
g(t, X_t) - g(0, X_0) = \int_0^t \sum_{i,j=1}^{d} \frac{\partial g(t, X_s)}{\partial x_i} \sigma_{ij}(X_s) dW_s^j
$$

This means that $g(t, X_t)$ is a local martingale under $\mathbb{Q}_i$, for $i = 1, 2$. Because g is bounded and continuous, by dominating convergence theorem we conclude that $g(t, X_t)$ is a true martingale. Therefore

$$
\begin{aligned}
\mathbb{E}_{\mathbb{Q}_1}[f(X_T)] &= \mathbb{E}_{\mathbb{Q}_1}[g(T, X_T)] \\
&= \mathbb{E}_{\mathbb{Q}_1}[g(0, X_0)] = g(T, x) \\
&= \mathbb{E}_{\mathbb{Q}_2}[g(0, X_0)] = \mathbb{E}_{\mathbb{Q}_2}[f(X_T)]
\end{aligned}
$$

The function f was chosen arbitrary in $C_0^\infty(\mathbb{R}^d)$ which is a class that determines probability measures, so $\mathbb{Q}_1 \circ X_T^{-1} = \mathbb{Q}_2 \circ X_T^{-1}$. By proposition 1.11, $\mathbb{P}_1^x = \mathbb{P}_2^x$.

We used the fact that the class $C_0^\infty(\mathbb{R}^d)$ determines probability measures*(problem 5.4.25 Karatzas & Shreve [1991]).*

Definition 3.2. A class of real functions $\mathcal{C}$ is called (measure)determining class, if any two measures μ_1,μ_2 coincide if

$$
\int f(x)\mu_1(dx) = \int f(x)\mu_2(dx) \qquad \forall f \in \mathcal{C}
$$

Remark 3.3. The class $C_0^\infty(\mathbb{R}^d)$ is measure determining class.
Proof: Suppose $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ is the Borel measurable space of $\mathbb{R}^d$. Suppose F is a closed set and let $\epsilon > 0$. We denote by $F^\epsilon = \{x \in \mathbb{R}^d : d(x, F) < \epsilon\}$ and choose

$$
f(x) = \max\{0, 1 - \frac{d(x, F)}{\epsilon}\}
$$

This function is bounded, uniformly continuous and absolutely integrable in $\mathbb{R}^d$. We next choose a function ϕ in $C_0^\infty(\mathbb{R}^d)$ such that $\int \phi(x)dx = 1$ and let $\phi_n(x) = \frac{\phi(x/n)}{n^d}$. We now consider the convolution $f_n := f * \phi_n$.
First we notice that f_n is an element of $C_0^\infty(\mathbb{R}^d)$. Because ϕ_n has compact support, there is a compact set K such that $\phi_n(x) = 0$, $\forall x \in \mathbb{R}^d/K$. This implies that the closed support of f_n is a subset of K, therefore bounded. By definition, the closed support of a function is a closed set, meaning f_n has compact support.
Fix $x_0 \in \mathbb{R}^d$. Because ϕ_n is continuously differentiable we get

$$f_n'(x_0) = \int f(x_0 - y)\phi_n'(y)dy = f * \phi_n'(x_0)$$

taking $|x - x_0| < \delta$, we get

$$\begin{aligned}
|f_n'(x) - f_n'(x_0)| &\leq \int |f(x-y) - f(x_0 - y)||\phi_n'(y)|dy \\
&\leq \int \frac{d(x, x_0)}{\epsilon}|\phi_n'(y)|dy \leq \frac{\delta}{\epsilon}\int_K |\phi_n'(y)|dy \\
&\leq \frac{\delta}{\epsilon}M
\end{aligned}$$

The second inequality is due to $|f(x) - f(y)| \leq \frac{d(x,y)}{\epsilon}$ and the last is by the continuity of ϕ_n' on compact. This means f_n is continuously differentiable and using the same argument and induction we conclude $f_n \in C_0^\infty$.
Next we want to use the fact that $f_n \to f$.

$$\begin{aligned}
|f_n(x) - f(x)| &= |\int \{f(x-y) - f(x)\}\phi_n(y)dy| \\
&\leq \int |f(x-y) - f(x)|\phi(y/n)/n^d dy \leq \int \frac{d(y,0)}{\epsilon}\frac{M'}{n^d}dy
\end{aligned}$$

Letting $n \to \infty$ we get $||f_n - f|| \to 0$.
Suppose now that for two measures μ and ν the following identity holds

$$\int h(x)\mu(dx) = \int h(x)\nu(dx) \qquad \forall h \in C_0^\infty(\mathbb{R}^d)$$

Choosing h to be the convolution $f * \phi_n$ we get by dominating convergence theorem

$$\int f(x)\mu(dx) = \int f(x)\nu(dx)$$

f is an approximation of the indicator function of F, satisfying

$$\mathbb{1}_F \leq f \leq \mathbb{1}_{F^\epsilon}$$

taking integrals

$$\mu(F) \leq \int f(x)\mu(dx) = \int f(x)\nu(dx) \leq \nu(F^{\epsilon})$$

the above relation holds for every $\epsilon > 0$, choosing $\epsilon_n = \epsilon/n$ and taking $n \to \infty$ we get $\mu(F) \leq \nu(F)$. By symmetry we conclude $\mu(F) = \nu(F)$, for every closed set on $\mathbb{R}^d$. Closed sets generates a Dynkin class equals to the Borel sets, implying $\mu = \nu$.

Solutions to the Martingale Problem conditioning that the process has continuous trajectories, are diffusion processes. We now study what properties will have such process if the MP has unique solution.

3.2 Strong Markov Property

Suppose that the SDE for b, σ is well posed. For each $x \in \mathbb{R}^d$ we let $\mathbb{P}^x$ to denote the unique probability law of this unique solution. The canonical process X equipped with $\{\mathbb{P}^x\}$ is a Strong Markov process*(Proposition I.5.1, R.Bass[1997]).*

Proposition 3.4. Suppose the martingale problem for $\mathcal{L}$ is well posed and $\mathbb{P}^x$ is the solution of the martingale problem for $\mathcal{L}$ with initial condition x, for each $x \in \mathbb{R}^d$. Then the process $(X_t, \mathbb{P}^x)$ is a Strong Markov process for every $x \in \mathbb{R}^d$.

Proof: Fix $x \in \mathbb{R}^d$. If the operator $\mathcal{L}$ is uniformly elliptic, then by Theorem 2.1 the unique probability measure $\mathbb{P}^x$ that solves the martingale problem for $\mathcal{L}$ and initial condition x, is the unique probability law of the weak solution to SDE for b, σ and x. By proposition I.5.1(R.Bass[1997]), the process $(X_t, \mathbb{P}^x)$ is a Strong Markov process.
In case $\mathcal{L}$ can degenerate, we consider the unique solution $\mathbb{P}^x$ of the martingale problem to be the probability measure induced by the process X from the extended probability space $(E, \mathcal{G}, \mathbb{Q})$ where the solution to SDE exists, i.e. $\mathbb{P}^x = \mathbb{Q} \circ X^{-1}$. By the uniqueness assumption, every two solutions to SDE, X^1, X^2 on $(E, \mathcal{G}, \mathbb{Q})$, will induce the same measure on $(\Omega, \mathcal{F})$. This implies that X^1 and X^2 have the same family of finite dimensional distributions and weak uniqueness holds and thus the process $(X_t, \mathbb{P}^x)$ forms a Strong Markov process.

3.3 Convergence results

Lemma 3.5. Suppose for every $n \in \mathbb{N}$, $\mathbb{P}^n$ is a probability measure on Ω. Let X_n be a sequence of continuous progressively measurable processes such that

$$\sup_n \sup_{t,\omega} |X_n(t,\omega)| < \infty$$

and each process X_n is a martingale under $\mathbb{P}^n$. Assume that there is a process X jointly continuous in t and ω progressively measurable, and a probability measure $\mathbb{P}$ on Ω such that $\mathbb{P}^n$ converges weakly to $\mathbb{P}$ and X_n converges uniformly on compact subsets of $\Omega \times [0,\infty)$ in X. Then X is a local martingale under $\mathbb{P}$.

Proof: Let $0 \leq s < t$ and f be a bounded continuous $\mathcal{F}_s$- measurable random variable. Then $X_s f$ is $\mathcal{F}_s$- measurable and $X_t f$ is $\mathcal{F}_t$- measurable r.v. Both $\mathbb{P} - a.s.$ are continuous. We consider the stopping time $\tau_M := \inf\{t \geq 0 : |X_t| \geq M\}$, because X is continuous $\tau_M \to \infty$ and the random variables $X_{s\wedge\tau_M} f$ $X_{t\wedge\tau_M} f$ are both bounded. Therefore by the assumption of weak convergence

$$\mathbb{E}_{\mathbb{P}}[X_{t\wedge\tau_M} f] = \lim_n \mathbb{E}_{\mathbb{P}^n}[X_{t\wedge\tau_M} f]$$

Moreover

$$\left|\mathbb{E}_{\mathbb{P}^n}[X_{t\wedge\tau_M} f] - \mathbb{E}_{\mathbb{P}^n}[X^n_{t\wedge\tau_M f}]\right| \leq ||f||\mathbb{E}_{\mathbb{P}^n}\left[|X_{t\wedge\tau_M} - X^n_{t\wedge\tau_M}|\right]$$

Because $X^n \to X$ uniformly on compacts, we have

$$\lim_n \sup_{s\leq t} \left|X^n_{s\wedge\tau_M} - X_{s\wedge\tau_M}\right| = 0$$

and X^n is uniformly bounded on compacts meaning also

$$\mathbb{E}_{\mathbb{P}^n}[\sup_n \sup_{s\leq t} |X^n_{s\wedge\tau_M}|] < \infty$$

For every t the sequence $|X^n_{t\wedge\tau_M}|$ is dominated by $\sup_n \sup_{t,\omega} |X^n_{t\wedge\tau_M}(\omega)| < \infty$ implying that the sequence $X^n_{t\wedge\tau_M}$ is uniformly integrable on compacts. By Vitali's theorem this means $\mathbb{E}[|X^n_{t\wedge\tau_M} - X_{t\wedge\tau_M}|] \to 0$ as $n \to \infty$.
Thus,

$$\mathbb{E}_{\mathbb{P}}[X_{t\wedge\tau_M} f] = \lim_n \mathbb{E}_{\mathbb{P}^n}[X^n_{t\wedge\tau_M} f] = \lim_n \mathbb{E}_{\mathbb{P}_n}[X^n_{s\wedge\tau_M} f] = \mathbb{E}_{\mathbb{P}}[X_{s\wedge\tau_M}]$$

Since X is a continuous process, the stopping time $\tau_M \to \infty$ a.s. as $M \to \infty$ and therefore X is a local martingale under $\mathbb{P}$.

We again consider the d-dimensional canonical process in its usual probability space and the usual second order linear differential operator $\mathcal{L}$ with diffusion and drift coefficients a and b respectively. A consequence of well posedness of the martingale problem for $\mathcal{L}$ is that under conditions of the coefficients if we have convergence of operators $\mathcal{L}_n$ to $\mathcal{L}$, there is stability of the solutions of the associated martingale problems *(Theorem 11.1.4. Stroock & Varadhan [1979])*. We will need theorem *7.3* from *Billingsley [1999]*.

Theorem 3.6. The sequence of probability measure $\{\mathbb{P}^n\}_n$ is tight if and only if the following conditions hold
(i) for all $\eta > 0$, there exist a positive number ζ and an $n_0 \in \mathbb{N}$, such that

$$\mathbb{P}^n(|X_0| \geq \zeta) \leq \eta, \qquad \forall n \geq n_0$$

(ii) for all $\epsilon > 0$,

$$\lim_{\delta \to 0} \limsup_n \mathbb{P}^n(w_x(\delta) \geq \epsilon) = 0$$

Proof : ($\Longrightarrow$) Suppose $\{\mathbb{P}^n\}_n$ is tight. Given $\eta > 0$, by definition we can choose a compact set K such that $\mathbb{P}^n(K) > 1-\eta$ for all n. By Arzelá - Ascoli theorem the set K is a subset of $\{\omega : |X_0(\omega)| \leq \zeta\}$ for some ζ and is also a subset of the event $\{\omega : \sup_{|t-s|\leq\delta} |X_t(\omega) - X_s(\omega)| \leq \epsilon\}$ for some δ. This implies that

$$\mathbb{P}^n(|X_0| > \zeta) \leq \eta$$

and

$$\mathbb{P}^n(w_x(\delta) > \epsilon) \leq \eta$$

for all n.
($\Longleftarrow$) We assume now, that the sequence of probability measures satisfy conditions (i) and (ii). Every probability measure on $C[0,\infty)^d$ is tight, therefore if we take finitely many probability measures $\{\mathbb{P}^n\}_{n=1}^{n_0}$, for every $\eta > 0, \epsilon > 0$ we can find ζ_n and δ_n, $1 \leq n \leq m$ such that, $\forall n$

$$\mathbb{P}^n(|X_0| > \zeta_n) \leq \eta$$

and

$$\mathbb{P}^n(w_x(\delta_n) > \epsilon) \leq \eta$$

then taking $\zeta = \max_{1\leq n\leq m}\{\zeta_n\}$ and $\delta = \min_{1\leq n\leq m}\{\delta_n\}$ we have conditions (i) and (ii) for finitely n's. Therefore it is sufficient to prove that conditions (i)-(ii)

implying tightness assuming $n_0 = 1$.
Given $\eta > 0$ and $\epsilon > 0$, choose $\zeta > 0$ and δ_k such that

$$\mathbb{P}^n(|X_0| > \zeta) \leq \eta \qquad \forall n$$

and

$$\mathbb{P}^n(w_x(\delta_k) > \epsilon) \leq \frac{\eta}{2^k} \qquad \forall n$$

we call $B := \{\omega : |X_0(\omega)| \leq a\}$ and $B_k := \{\omega : w_{x(\omega)}(\delta_k) \leq \frac{1}{k}\}$ and we have $\mathbb{P}^n(B) \geq 1 - \eta$ and $\mathbb{P}^n(B_k) \geq 1 - \frac{\eta}{2^k}$. Finally we define the event

$$A = B \cap \bigcap_k B_k$$

and get

$$\mathbb{P}^n\left(\overline{A}\right) \geq 1 - 2\eta \qquad \forall n$$

The set A satisfies the conditions of Azarlá- Ascoli meaning it is relatively compact. This implies that $\overline{A}$ is compact, meaning the sequence $\mathbb{P}^n$ is tight.

Theorem 3.7. Suppose that a and b are continuous uniformly bounded and a is positive semi definite $d \times d$ function matrix. Suppose the martingale problem for $\mathcal{L}$ is well posed and $\mathbb{P}^x$ is the unique solution for $\mathrm{MP}(\mathcal{L}, \delta_x)$, for each $x \in \mathbb{R}^d$. Next assume that there are sequences $\{a_n\}$ and $\{b_n\}$ of measurable functions such that a_n is positive semi definite for each $n \geq 1$ and that for all $T > 0$ and $R > 0$ the following hold

$$\sup_n \sup_{|x| \leq R} ||a_n(x)|| + |b_n(x)| < \infty$$

and

$$\lim_n \int_0^T \sup_{|x| \leq R} \big(||a_n(x) - a(x)|| + |b_n(x) - b(x)|\big) ds = 0$$

If $\mathbb{P}^n$ be a solution to the $\mathrm{MP}(\mathcal{L}_n, \delta_{x_n})$ and $x_n \to x$, then $\mathbb{P}^n \Rightarrow \mathbb{P}$

Proof:
First we show that the sequence $\mathbb{P}^n$ is relatively compact, i.e. that there is a convergent subsequence $\mathbb{P}^{n_k}$. Each $\mathbb{P}^n$ is a probability measure on $C[0, \infty)^d$. We are going to use the Arzelá-Ascoli type theorem(7.3) from Billingsley([1999]) to establish tightness, then by Prohorov's theorem we will conclude that the sequence is relatively compact. The first condition holds for $\zeta = \sup_n |x_n| + 1$. For the second condition we need to estimate the

modulus of continuity of the paths of the cannonical process. Fix $0 < s \leq 1$ and let Λ be the bound of b, if $\epsilon \geq 2\Lambda(t-s)$, from remark 3.8. we get

$$\mathbb{P}^n\big(\sup_{s\leq t\leq s+\delta}|X_t - X_s| \geq \epsilon\big) \leq \mathbb{P}^n\big(\sup_{t\leq s+\delta}|M^s_t| \geq \epsilon/2\big) \leq 2e^{-\frac{\epsilon^2}{8c\delta}}$$

This bound does not depend on n, therefore we conclude for every $\epsilon > 0$

$$\lim_{\delta\to 0}\limsup_n \mathbb{P}^n\big(w_x(\delta) \geq \epsilon\big) = 0$$

Now we have that the family $\mathbb{P}^n$ is sequentially relatively compact, so there is a subsequence $\mathbb{P}^{n_k}$ converges to a measure in $(C[0,\infty)^d, \mathcal{B}(C[0,\infty)^d))$ under the weak topology. We consider Q to be a limit point of a convergent subsequence $\mathbb{P}^{n_k}$ of $\mathbb{P}^n$. We are going to show that Q is a solution to the local martingale problem for $\mathcal{L}$ with initial condition x.
By hypothesis $\mathbb{P}^{n_k}$ is a solution to $\mathrm{MP}(\mathcal{L}_{n_k}, \delta_{x_{n_k}})$, X_0 is a continuous and bounded random variable under $\mathbb{P}^{n_k}$ and $x_n \to x$, so we get

$$\mathbb{P}^{n_k}(X_0 = x_{n_k}) = 1$$

and

$$\mathbb{E}_{\mathbb{P}^{n_k}}[h(X_0)] \to \mathbb{E}_Q[h(X_0)]$$

for every bounded continuous function h hence, $Q(X_0 = x) = 1$.
Next we have to establish the martingale property. We consider a bounded function f in $C^2(\mathbb{R}^d)$ with bounded first and second partial derivatives and the exit time $\tau_R = \inf\{t \geq 0 : |X_t| \geq R\}$. From our limit assumption we get that $\int_0^T \mathcal{L}_n f(X_r)dr$ converges uniformly on compacts to $\int_0^T \mathcal{L}f(X_r)dr$. We define

$$M^n_t := f(X_t) - f(X_0) - \int_0^t \mathcal{L}_n f(X_r)dr$$

and

$$M_t := f(X_t) - f(X_0) - \int_0^t \mathcal{L}f(X_r)dr$$

So we have that $M^{n_k}_{t\wedge\tau_R}$ converges uniformly on compacts to $M_{t\wedge\tau_R}$ and that

$$\sup_k \sup_{t,\omega} |M^{n_k}_{t\wedge\tau_R}| < \infty$$

By Lemma 3.5 we get that the process $M_{t\wedge\tau_R}$ is a martingale under Q. By the well-posedness of the martingale problem for $\mathcal{L}$ we conclude that Q equals $\mathbb{P}^x$ on $\mathcal{F}_{\tau_R}$, for every $R > 0$. If f is a function in $C^2(\mathbb{R}^d)$ but not necessarily bounded we consider a bounded function f_R with bounded first and second partial derivatives that equals f on the ball $B(0,R)$. Then the process

$$M_t^{f_R} = f_R(X_t) - f_R(X_0) - \int_0^t \mathcal{L} f_R(X_r)dr$$

is a martingale under Q, which implies that for arbitrary $f \in C^2(\mathbb{R}^d)$, the process $M_{t\wedge\tau_R}$ is a martingale under Q.
We have till now, that every subsequental limit of $\mathbb{P}^n$ coincides with $\mathbb{P}^x$ on the events of $\mathcal{F}_{\tau_R}$, for every $R > 0$. Finally we consider g to be a bounded and continuous function, then

$$\left|\mathbb{E}_{\mathbb{P}_{n_k}}[g(X_t)] - \mathbb{E}_{\mathbb{P}}[g(X_t)]\right| \leq$$

$$\left|\mathbb{E}_{\mathbb{P}_{n_k}}[g(X_t)\mathbb{1}_{\tau_R \leq t}] - \mathbb{E}_{\mathbb{P}^x}[g(X_t)\mathbb{1}_{\tau_R\leq t}] + \mathbb{E}_{\mathbb{P}^{n_k}}[g(X_t)\mathbb{1}_{\tau_R>t}] - \mathbb{E}_{\mathbb{P}^x}[g(X_t)\mathbb{1}_{\tau_R>t}]\right|$$
$$\leq ||g||\mathbb{P}^{n_k}(\tau_R \leq t) + ||g||\mathbb{P}^x(\tau_R \leq T) + \left|\mathbb{E}_{\mathbb{P}^{n_k}}[g(X_t)\mathbb{1}_{\tau_R>t}] - \mathbb{E}_{\mathbb{P}^x}[g(X_t)\mathbb{1}_{\tau_R>t}]\right|$$

letting $k \to \infty$, because the event $\{\tau_R \leq t\}$ is closed in Ω we get

$$\left|\mathbb{E}_Q[g(X_t)] - \mathbb{E}_{\mathbb{P}^x}[g(X_t)]\right| \leq 2||g||\mathbb{P}^x(\tau_R \leq t)$$

letting $R \to \infty$, we get $\mathbb{P}^x(\tau_R \leq t) \to 0$, implying

$$\mathbb{E}_Q[g(X_t)] = \lim_k \mathbb{E}_{\mathbb{P}^{n_k}}[g(X_t)] = \mathbb{E}_{\mathbb{P}^x}[g(X_t)]$$

for every bounded and continuous function g. This means that any convergent subsequence has limit point equal to $\mathbb{P}^x$ and this means that the whole sequence converges to $\mathbb{P}^x$ *(Theorem 2.6. Billingsley[1999]).*

The next remark is *exercise I.8.13 from Bass [1995]*, we imitate the proof of *Proposition I.4.8 (Bass[1995])*, which is the result in the case of Brownian motion.

Remark 3.8. If M is a continuous martingale with $M_0 = 0$, then

$$\mathbb{P}\Big(\sup_{s\leq t}|M_s| > \eta, \langle M\rangle_t < \theta\Big) \leq 2e^{-\frac{\eta^2}{2\theta}}$$

Proof: Let $\gamma > 0$ and let $N_t := exp\left(\gamma M_t - \frac{1}{2}\gamma^2\langle M\rangle_t\right)$, then if $\mathbb{E}[e^{\frac{1}{2}\langle M\rangle_t}] < \infty$ for every t, N is a continuous martingale and by Doob's inequality

$$\mathbb{P}(\sup_{s\leq t} M_s > \eta, \langle M\rangle_t < \theta) = \mathbb{P}(\sup_{s\leq t} N_s > e^{\gamma\eta - \frac{1}{2}\gamma^2\theta}) \leq e^{-\gamma\eta + \frac{1}{2}\gamma^2\theta}\mathbb{E}[N_t]$$

We choose $\gamma = \frac{\eta}{\theta}$, and get

$$\mathbb{P}(\sup_{s \leq t} M_s > \eta, \langle M \rangle_t < \theta) \leq e^{-\frac{\eta^2}{2\theta}}$$

We repeat the argument for $-M$ and get

$$\mathbb{P}(\sup_{s \leq t}(-M_s) > \eta, \langle M \rangle_t < \theta) \leq e^{-\frac{\eta^2}{2\theta}}$$

we then, add the two inequalities.

A consequence of convergence of probability measures is the following existence result which is the Skorohod theorem [1965].

Theorem 3.9. Suppose a and b are uniformly bounded continuous functions and a is positive semi definite. Then for each $x \in \mathbb{R}^d$, the MP$(\mathcal{L}, \delta_x)$ admits a solution.

Proof: We consider the associated stochastic differential equation $SDE(b, \sigma, x)$ where $a = \sigma\sigma^T$. We choose sequences of $C^2(\mathbb{R}^d)$ uniformly bounded functions σ_{ij}^n and b_i^n that converge uniformly on compacts to σ_{ij} and b_i respectively (this approximation can be done due to Weierstrass approximation theorem). Next consider

$$dX_t^n = b^n(X_t^n)dt + \sigma^n(X_t^n)dW_t$$

Because the coefficients are Lipschitz continuous, the SDE(b^n, σ^n, x) has a unique weak solution. By theorem 2.1. there is a unique probability measure $\mathbb{P}^n$ solving MP$(\mathcal{L}_n, \delta_x)$.
As in theorem 3.7. the sequence $\{\mathbb{P}^n\}_n$ is relatively compact, therefore has a convergent subsequence $\mathbb{P}^{n_k}$. Let $\mathbb{P}$ be this limit point. This probability measure will be the solution to MP$(\mathcal{L}, \delta_x)$. Every diffusion X^n starts from x and $\lim_k \mathbb{P}^{n_k} = \mathbb{P}$, so

$$\mathbb{P}(X_0 = x) = 1$$

Let $f \in C^2(\mathbb{R}^d)$ and $f_R \in C^2(\mathbb{R}^d)$ a bounded continuous function with bounded partial first and second derivatives such that $f_R = f$ on $B(0, R)$. Because

$$\sup_n \sup_{t,\omega} |M_{t \wedge \tau R}^f(\omega)| < \infty$$

where $\tau_R = \inf\{t \geq 0 : |X_t| \geq R\}$ and by Lemma 3.5. the stopped process $M_{t \wedge \tau_R}^f$ is a martingale under $\mathbb{P}$. Because $\tau_R \to \infty$ as $R \to \infty$ we conclude that M_t^f is a local martingale under $\mathbb{P}$.

In case of uniform ellipticity we can eliminate the conditions for the drift coefficient in order to prove existence and uniqueness of the martingale problem*(Theorem 6.4.3. Stroock & Varadhan [1979], Theorem VI.3.1. Bass [1997])*.

Theorem 3.10. Suppose $\mathcal{L}$ is uniformly elliptic with bounded measurable coefficients a and b. Define

$$\mathcal{L}'f(x) = \frac{1}{2}\sum_{i,j=1}^{d} a_{ij}(x)\partial_{ij}f(x)$$

If the martingale problem for $\mathcal{L}'$ is well posed, then the martingale problem for $\mathcal{L}$ is well posed. In case of well posedness, if the family of solutions of MP for $\mathcal{L}'$ is measurable, then the family of solutions of MP for $\mathcal{L}$ is measurable.

Proof : Suppose that the MP$(\mathcal{L}, \delta_x)$ has a unique solution. Let $\mathbb{P}_1$ and $\mathbb{P}_2$ are two solutions to MP$(\mathcal{L}, \delta_x)$. The test process M_t^f is a martingale under $\mathbb{P}_i$, $i = 1, 2$. We define two new probability measures Q_1 and Q_2 such that on the events of $\mathcal{F}_t$, the density functions $\frac{dQ_1}{s\mathbb{P}_1}$ and $\frac{dQ_2}{d\mathbb{P}_2}$ are both

$$N_t := \exp\Big(- \int_0^t ba^{-1}(X_r)dX_r - \frac{1}{2}\int_0^t ba^{-1}b^T(X_r)dr\Big)$$

The process N_t is the stochastic exponential of a martingale process therefore an $\mathcal{F}_t$- martingale under $\mathbb{P}_1$ and $\mathbb{P}_2$.

From Theorem 2.1. we know that X_t is a weak solution to

$$dX_t = b(X_t)dt + \sigma(X_t)dW_t$$

for some Brownian motion W_t. This means that $d\langle X_\cdot\rangle_t = a(X_t)dt$ and so by Itô's formula

$$df(X_t) = \mathcal{L}f(X_t)dt + \nabla f(X_t)\sigma(X_t)dW_t$$

implying

$$dM_t^f = \nabla f(X_t)d\Big(X_t - \int_0^t b(X_r)dr\Big)$$

therefore

$$\Big\langle \int_0^\cdot ba^{-1}(X_r)dX_r, M^f_\cdot \Big\rangle_t = \int_0^t ba^{-1}(X_r)\nabla f(X_r)d\langle X_\cdot\rangle_r$$

$$= \int_0^t \sum_{i,j=1}^d (ba^{-1})_j(X_r)\partial_j f(X_r)d\langle X^i_\cdot, X^j_\cdot\rangle_r$$

$$= \int_0^t \sum_{i=1}^d b_i(X_r)\partial_i f(X_r)dr$$

By Girsanov's theorem, the process

$$M_t - \Big\langle -\int_0^\cdot ba^{-1}(X_r)dX_r, M^f_\cdot \Big\rangle_t = f(X_t) - f(X_0) - \int_0^t \mathcal{L}' f(X_r)dr$$

is a martingale under Q_1 and Q_2 with the same quadratic variation as under $\mathbb{P}_i$, $i = 1, 2$. Because

$$Q_i(X_0 = x) = \mathbb{P}_i(X_0 = x) = 1 \qquad i = 1, 2$$

by the uniqueness of the martingale problem for $\mathcal{L}'$, $Q_1 = Q_2$. This means

$$\mathbb{P}_1(A) = \mathbb{E}_{Q_1}[N_t\mathbb{1}_A] = \mathbb{E}_{Q_2}[N_t\mathbb{1}_A] = \mathbb{P}_2(A)$$

for every $A \in \mathcal{F}_t$ and for each t.

Finally suppose that the mapping $x \to Q^x(F)$ is measurable for every $F \in \mathcal{F}$. If $A \in \mathcal{F}_t$, then

$$\mathbb{P}^x(A) = \mathbb{E}_{Q^x}\Big[\exp\Big(\int_0^t (ba^{-1}(X_r)dX_r + \int_0^t (ba^{-1}b)(X_r)dr\Big)\mathbb{1}_A\Big]$$

Suppose $x_n \to x$ and F is a closed set of Ω. We define $A := F - x$ and $A_n := F - x_n$. By Fatou's lemma we get

$$\limsup \mathbb{P}^{x_n}(F) = \limsup \mathbb{E}_{Q^0}\Big[\frac{1}{N_t}\mathbb{1}_{A_n}\Big] \leq \mathbb{E}_{Q^0}\Big[\limsup \frac{1}{N_t}\mathbb{1}_{A_n}\Big] = \mathbb{P}^x(F)$$

and thus $\mathbb{P}^x(F)$ is upper semi continuous and therefore measurable, for every closed set F. This is a Dynkin class that generates the Borel and so we have the desired property for $\mathcal{F}$ which is a sub-σ algebra of the Borel.

As a matter of fact, in case of uniformly elliptic operators, existence for the associated martingale problem holds under weaker conditions of both coefficients, specifically we can eliminate any smoothness assumption *(Theorem VI.1.3. Bass [1997])*.

Theorem 3.11. Suppose $\mathcal{L}$ is uniformly elliptic where $a : \mathbb{R}^d \to S^d$ and $b : \mathbb{R}^d \to \mathbb{R}^d$ are bounded measurable functions. Then for each $x \in \mathbb{R}^d$, there exists a solution to MP$(\mathcal{L}, \delta_x)$.

Proof: By theorem 3.10, it is sufficient to prove that the MP$(\mathcal{L}', \delta_x)$ has a solution, where

$$\mathcal{L}' f(x) = \frac{1}{2} \sum_{i,j=1}^{d} a_{ij}(x) \partial_{ij} f(x)$$

We consider sequences a_{ij}^n for every $i, j = 1, ..., d$ such that a_{ij}^n are elements of $C^2(\mathbb{R}^d)$ and $a_{ij}^n \to a_{ij}$ almost everywhere. Then the corresponding to a^n operator $\mathcal{L}'_n$ is uniformly elliptic and the associated martingale problems admit a unique solution when starting from $x \in \mathbb{R}^d$. Let $\mathbb{P}^n \in$ MP$(\mathcal{L}'_n, \delta_x)$, similarly with previous results we will prove that $\{\mathbb{P}^n\}$ has a limit point and that this is a solution to MP$(\mathcal{L}', \delta_x)$.
As in Theorem 3.7 the sequence $\{\mathbb{P}^n\}$ is relatively compact. If $\mathbb{P}^{n_k}$ is a a convergent subsequence of $\mathbb{P}^n$, we consider a further subsequence $\mathbb{P}^{n_{k_m}}$ such that $a^{n_{k_m}} \to a$ in $\mathbb{L}_p$. We need to show, for every $A \in \mathcal{F}_s$ and for $f \in C^2(\mathbb{R}^d)$

$$\lim_m \mathbb{E}_{\mathbb{P}^{n_{k_m}}} \left[\{M_t^f - M_s^f\} \mathbb{1}_A \right] = \mathbb{E}_{\mathbb{P}} \left[\{M_t^f - M_s^f\} \mathbb{1}_A \right]$$

Let g be bounded and continuous function on $\mathbb{R}^d$, using the same localization argument as in theorem 3.7. equivalently we need to show.

$$\mathbb{E}_{\mathbb{P}^{n_{k_m}}} \Big[g(X_s) \int_s^t \mathcal{L}'_{n_k} f(X_r) dr \Big] \underset{m}{\to} \mathbb{E}_{\mathbb{P}} \Big[g(X_s) \int_s^t \mathcal{L}' f(X_r) dr \Big]$$

for every bounded function f in $C^2(\mathbb{R}^d)$ with bounded first and second derivatives.
Let m_0 in $\mathbb{N}$.

$$\Big| \mathbb{E}_{\mathbb{P}^{n_{k_m}}} \Big[g(X_s) \int_s^t \mathcal{L}'_{n_{k_m}} f(X_r) dr \Big] - \mathbb{E}_{\mathbb{P}} \Big[g(X_s) \int_s^t \mathcal{L}' f(X_r) dr \Big] \Big| \leq$$

$$\Big| \mathbb{E}_{\mathbb{P}^{n_{k_m}}} [g(X_s) \int_s^t \mathcal{L}'_{n_{k_m}} f(X_r) - \mathcal{L}'_{n_{k_{m_0}}} f(X_r) dr] \Big| +$$

$$\Big| \mathbb{E}_{\mathbb{P}^{n_{k_m}}} \Big[g(X_s) \int_s^t \mathcal{L}'_{n_{k_{m_0}}} f(X_r) dr \Big] - \mathbb{E}_{\mathbb{P}} \Big[g(X_s) \int_s^t \mathcal{L}' f(X_r) dr \Big] \Big| \leq$$

$$\Big| \mathbb{E}_{\mathbb{P}^{n_{k_m}}} \Big[g(X_s) \int_s^t \mathcal{L}'_{n_{k_m}} f(X_r) - \mathcal{L}'_{n_{k_{m_0}}} f(X_r) dr \Big] \Big| + \qquad (i)$$

$$\Big|\mathbb{E}_{\mathbb{P}^{n_{k_m}}}\Big[g(X_s)\int_s^t \mathcal{L}'_{n_{k_{m_0}}} f(X_r)dr\Big] - \mathbb{E}_{\mathbb{P}}\Big[g(X_s)\int_s^t \mathcal{L}'_{n_{k_{m_0}}} f(X_r)dr\Big] + \qquad (ii)$$

$$\Big|\mathbb{E}_{\mathbb{P}}\Big[g(X_s)\int_s^t \mathcal{L}'_{n_{k_{m_0}}} f(X_r) - \mathcal{L}' f(X_r)dr\Big]\Big| \qquad (iii)$$

quantity (ii) tends to zero bythe weak convergence of $\mathbb{P}^{n_{k_m}}$. For (i) and (iii) we use Alexandrov's estimate

$$\Big|\mathbb{E}_{\mathbb{P}}\Big[g(X_s)\int_s^t \mathcal{L}'_{n_{k_{m_0}}} f(X_r) - \mathcal{L}' f(X_r)dr\Big]\Big| \leq C_1 ||\mathcal{L}'_{n_{k_{m_0}}} - \mathcal{L}'||_{\mathbb{L}_p}$$

and

$$\Big|\mathbb{E}_{\mathbb{P}^{n_{k_m}}}\Big[g(X_s)\int_s^t \mathcal{L}'_{n_{k_m}} f(X_r) - \mathcal{L}'_{n_{k_{m_0}}} f(X_r)dr\Big]\Big| \leq C_2 ||\mathcal{L}'_{n_{k_{m_0}}} - \mathcal{L}'||_{\mathbb{L}_p}$$

while C_1 and C_2 depend only on the elliptic bound of a and p, s, t. Finally

$$||\mathcal{L}'_{n_{k_{m_0}}} - \mathcal{L}'||_{\mathbb{L}_p} \leq M\Big(\int_s^t ||a^{n_{k_m}}(x) - a(x)||^p dx\Big)^{1/p}$$

and

$$||\mathcal{L}'_{n_{k_m}} - \mathcal{L}'_{n_{k_{m_0}}}||_{\mathbb{L}_p} \leq M\Big(\int_s^t ||a^{n_{k_m}}(x) - a^{n_{k_{m_0}}}(x)||^p dx\Big)^{1/p}$$

while M depends only on the bounds of the second partial derivatives of f. Let $m \to \infty$ and $m_0 \to \infty$, by the $\mathbb{L}_p$ convergence of $a^{n_{k_m}}$ to a, we have the result.

4 Application

An interesting application of uniqueness of martingale problems is convergence of Markov chains to a diffusion. In this section, we will consider a discretization of our cannonical process, which under a probability measure will be a time-homogenous Markov chain starting from x in $\mathbb{R}^d$ with some transition function. Then we'll see that if the martingale problem for $\mathcal{L}$ is well posed, under conditions of the coefficients and of the transition function, for each initial condition x, the probability laws of the Markov chains will converge weakly to the unique solution of MP *(Chapter 11.2[p.266-272], Stroock & Varadhan [1997])*.

Let $h > 0$, $x \in \mathbb{R}^d$ and $\Pi_h(x, \cdot)$ a transition probability function. We consider a probability measure $\mathbb{P}_h^x$ on Ω, with the following properties
(i)

$$\mathbb{P}_h^x(X_0 = x) = 1$$

(ii) for every $k \geq 0$

$$\mathbb{P}_h^x\Big(X_t = \frac{(k+1)h - t}{h} X_{kh} + \frac{t - kh}{h} X_{(k+1)h}, \qquad kh \leq t < (k+1)h\Big) = 1$$

(iii) for every $\Gamma \in \mathbb{B}(\mathbb{R}^d)$

$$\mathbb{P}_h^x(X_{(k+1)h} \in \Gamma | \mathcal{F}_{kh}) = \Pi_h(X_{kh}, \Gamma)$$

The stochastic process $\{X_{kh}\}_{k=0}^{\infty}$ is a process on the space of real sequences. For each $k \in \mathbb{N}_0$, the random variable X_{kh} sends each ω to it's kh^{th} coordinate. The third property says that the discrete time process $\{X_{kh}\}_{k=0}^{\infty}$ is Markov chain with transition function $\Pi_h(x, \cdot)$ starting from x. We will need the definition of the discrete parameter martingale problem.

Let temporarily Ω be the space or real sequences and X be the canonical process on $(\Omega, \mathcal{F})$ where $\mathcal{F}$ and the $\mathcal{F}_n$ are the usual σ- algebra and the usual filtration. Suppose $\Pi_n(x, \cdot)$ is a sequence of transition probability functions for X under some probability measure, we consider an operator A_n who acts on continuous and bounded functions of $\mathbb{R}^d$ in the following way

$$A_n f(x) = \int \{f(y) - f(x)\} \Pi_n(x, dy)$$

Definition 4.1. A probability measure $\mathbb{P}$ is a solution to the discrete martingale problem for A_n with initial condition x if

$$\mathbb{P}(X_0 = x) = 1$$

and

$$f(X_n) - f(X_0) - \sum_{i=0}^{n-1} A_i f(X_i)$$

is a martingale with respect to $\mathcal{F}_n$, $\forall f \in C_b(\mathbb{R}^d)$.

If the process X_n is a Markov chain with transition functions $\Pi_n(x,\cdot)$, under a probability measure $\mathbb{P}$, then the test process M_n^f is always a martingale under this probability measure *(Exercise 6.7.1 Stroock & Varadhan [1979]).*

Proposition 4.2. The stochastic process X_n is a Markov chain with transition probability function $\Pi_n(x,\cdot)$ if and only if M_n^f is a martingale for every $f \in C_b(\mathbb{R}^d)$, where

$$M_n^f = f(X_n) - \sum_{i=0}^{n-1} A_i f(X_i)$$

Proof : ($\Longrightarrow$) Because $A_n f(X_n) = \mathbb{E}[f(X_{n+1})|X_n] - f(X_n)$ and by the Markov property, we got

$$\begin{aligned}\mathbb{E}[M_{n+1}^f - M_n^f|\mathcal{F}_n] &= \mathbb{E}[f(X_{n+1}) - f(X_n) - A_n f(X_n)|\mathcal{F}_n] \\ &= \mathbb{E}[f(X_{n+1})|\mathcal{F}_n] - \mathbb{E}[f(X_{n+1})|X_n] = 0\end{aligned}$$

($\Longleftarrow$) Suppose M_n^f is a martingale process under a probability measure $\mathbb{P}$ for each $f \in C_b(\mathbb{R}^d)$. We have by definition

$$A_n f(x) = \int \{f(y) - f(x)\}\Pi_n(x, dy)$$

and

$$A_n f(X_n) = \int \{f(y) - f(X_n)\}\Pi_n(X_n, dy) = \mathbb{E}_{\Pi_n(X_n,\cdot)}[f(X_{n+1})|X_n] - f(X_n)$$

by the martingale property

$$\mathbb{E}_{\mathbb{P}}[f(X_{n+1})|\mathcal{F}_n] = \mathbb{E}_{\mathbb{P}}[\mathbb{E}_{\Pi_n}[f(X_{n+1})|X_n]|\mathcal{F}_n] = \mathbb{E}_{\Pi_n}[f(X_{n+1})|X_n]$$

the above equality means

$$\mathbb{P}(X_{n+1} \in A|\mathcal{F}_n) = \Pi_n(X_n, A)$$

which says that X is a Markov chain with transition function Π.

Given that the chain starts from a point $x \in \mathbb{R}^d$, there is a unique probability measure P under which X_n is a Markov chain with transition probability function $\Pi_n(x,\cdot)$. Instead of saying that X is Markov with transition function Π is equivalent to say that the martingale problem for A is well posed.

We are again in the space of continuous functions $\Omega = C[0,\infty)^d$. Suppose a and b are continuous uniformly bounded functions and a is positive semi definite. Let $h > 0$ and $\Pi_h(\cdot,\cdot)$ be a Markov kernel. Define functions a^h_{ij}, b^h_i and $\Delta^\epsilon_h(x)$ as follows

$$a^h_{ij}(x) = \int_{|y-x|\leq 1} (y_i - x_i)(y_j - x_j)\Pi_h(x, dy)$$

$$b^h_i(x) = \int_{|y-x|\leq 1} (y_i - x_i)\Pi_h(x, dy)$$

$$\Delta^\epsilon_h(x) = \frac{1}{h}\Pi_h(x, B(x,\epsilon)^c), \qquad \text{for} \quad \epsilon > 0$$

Assume that (i) a^h and b^h are uniformly bounded such that (ii) $a^h \to a$ and $b^h \to b$ uniformly on compacts as $h \to 0$ and that (iii)

$$\lim_{h\to 0} \sup_{x\in\mathbb{R}^d} \Delta^\epsilon_h(x) = 0, \qquad \text{for} \quad \epsilon > 0$$

Theorem 4.3. Suppose the martingale problem for $\mathcal{L}$ is well posed where the diffusion and drift coefficients are continuous uniformly bounded functions and a is positive semi definite. Suppose the above conditions (i)-(iii) hold. If $\mathbb{P}^{x_0}$ is the unique solution of MP$(\mathcal{L}, \delta_{x_0})$ and if $\mathbb{P}^{x_0}_h$ is the probability measure corresponding to the Markov chain with transition function $\Pi_h(x,\cdot)$ starting from $x_0 \in \mathbb{R}^d$, then $\mathbb{P}^{x_0}_h \Rightarrow \mathbb{P}^{x_0}$, as $h \to 0$.

Proof : As in previous convergence result, we have to show that the family $\{\mathbb{P}^{x_0}_h\}_{h\geq 0}$ of probability measures is relatively compact, then we'll see that any subsequental limit point solves the MP$(\mathcal{L}, \delta_{x_0})$ and by this and the well posedness of MP we will conclude the convergence of measures.

In order to prove compactness it is sufficient for $\{X_{kh}\}_{k\in\mathbb{N}}$ to satisfy the following conditions *(Theorem 1.4.11 Stroock & Varadhan [1979])*
1. For every $f \in C^\infty_0(\mathbb{R}^d)$, there is a constant C_f which does not depend on h such that $f(X_{kh}) + C_f kh$ is a submartingale under $\mathbb{P}^{x_0}_h$.
2. For every $\epsilon > 0$ and for every $T > 0$

$$\sum_{0\leq jh\leq T} \mathbb{P}^{x_0}_h(|X_{(k+1)h} - X_{kh}| \leq \epsilon) \underset{h\to 0}{\to} 0$$

Let $f \in C_0^\infty(\mathbb{R}^d)$.

$$\big|\frac{1}{h}A_h f(x)\big| =$$

$$\big|\frac{1}{h}\int_{|y-x|\le 1}\{f(y)-f(x)\}\Pi_h(x,dy) + \int_{|y-x|>1}\{f(y)-f(x)\}\Pi_h(x,dy)\big| \le$$

$$\frac{|\nabla f(x)|}{h}\int_{|y-x|\le 1}|y-x|\Pi_h(x,dy) + \frac{\sum_{i,j=1}^d ||\partial_{ij}f||}{2h}\int_{|y-x|\le 1}|y-x|^2\Pi_h(x,dy)$$

$$+2||f||\Delta_h^1(x)$$

implying

$$\sup_{h>0}\big|\frac{1}{h}A_h f(x)\big| \le \sup_{h>0}||b^h||\frac{|\nabla f(x)|}{h} + \sup_{h>0}\frac{\sum_{ij}||\partial_{ij}f||}{2h}||a^h|| + 2||f||\sup_{h>0}\Delta_h^1(x)$$

implying that there exists a constant C_f depending only on the function f such that

$$\sup_{h>0}\big|\frac{1}{h}A_h f(x)\big| \le C_f$$

therefore

$$\mathbb{E}_{\mathbb{P}_h^{x_0}}[f(X_{(k+1)h}) - f(X_{kh}) + C_f h|\mathcal{F}_{kh}] \ge \mathbb{E}_{\mathbb{P}_h^{x_0}}[M_{k+1}^f - M_k^f|\mathcal{F}_{kh}] = 0$$

and this means that $f(X_{kh}) + C_f kh$ is a submartingale under $\mathbb{P}_h^{x_0}$.
Let $\epsilon > 0$ and $T > 0$.

$$\mathbb{P}_h^{x_0}(|X_{(k+1)h} - X_{kh}| \ge \epsilon) = \mathbb{E}_{\mathbb{P}_h^{x_0}}[\mathbb{E}_{\mathbb{P}_h^{x_0}}\mathbb{1}_{\{X_{(k+1)}\in B(X_{kh},\epsilon)^c\}}|\mathcal{F}_{kh}]]$$

$$= \mathbb{E}_{\mathbb{P}_h^{x_0}}[\Pi_h(X_{kh}, B(X_{kh},\epsilon)^c)] \le \sup_{x\in\mathbb{R}^d} h\Delta_h^\epsilon(x) \underset{h\to 0}{\to} 0$$

therefore

$$\sum_{1\le jh\le T}\mathbb{P}_h^{x_0}(|X_{(j+1)h} - X_{jh}| \ge \epsilon) = \sum_{i\le T/h}\mathbb{P}_h^{x_0}(|X_{i+h} - X_i| \ge \epsilon)$$

$$\le \frac{T+1}{h}\sup_{x\in\mathbb{R}^d} h\Delta_h^\epsilon(x) \underset{h\to 0}{\to} 0$$

The two condition are fulfilled and the family $\{\mathbb{P}_h^{x_0}\}_{h>0}$ is relatively compact.

In this point suppose that

$$\frac{A_h f}{h} \to \mathcal{L}f$$

uniformly on compact sets of $\mathbb{R}^d$, for every $f \in C_0^\infty(\mathbb{R}^d)$.

Let $\{h_n\}_n$ be a non increasing sequence such that $h_n \to 0$, then the sequence of probability measures $\mathbb{P}_{h_n}^{x_0}$ has a convergent subsequence. Suppose $\mathbb{P}$ is the subsequental limit point, then $\mathbb{P}$ is a solution to the martingale problem for $\mathcal{L}$ with initial condition x_0 . First as in theorem 3.7 we have $\mathbb{P}(X_0 = x_0) = 1$. Next, let g be a bounded and continuous function and $s \leq t$. We have that $\mathbb{P}_{h_n}^{x_0}$ is the solution to the discrete parameter martingale problem for A_{h_n} starting from x_0. Set

$$k_n = [\frac{s}{h_n}] + 1 \qquad \text{and} \qquad l_n = [\frac{t}{h_n}] + 1$$

then

$$\mathbb{E}_{\mathbb{P}_{h_n}^{x_0}}[\{f(X_{l_n h_n}) - f(X_{k_n h_n}) - \sum_{j=k_n}^{l_n - 1} A_{h_n} f(X_{j h_n})\} g(X_s)] = 0$$

using the fact that $\frac{A_h f}{h} \to \mathcal{L}f$ uniformly on compacts we have

$$M_{l_n h_n}^h - M_{k_n h_n}^h := f(X_{l_n h_n}) - f(X_{k_n h_n}) - \sum_{r=k_n h_n}^{(l_n - 1)h_n} A_{h_n} f(X_r)$$

$$\to f(X_t) - f(X_s) - \int_s^t \mathcal{L}f(X_r) dr = M_t^f - M_s^f$$

uniformly on compacts. Because $M_t^f - M_s^f$ is bounded and continuous

$$\mathbb{E}_{\mathbb{P}_{h_n}^{x_0}}[\{M_{l_n h_n}^h - M_{k_n hn}^h\} g(X_s)] \to \mathbb{E}_{\mathbb{P}}[\{M_t^f - M_s^f\} g(X_s)]$$

meaning

$$\mathbb{E}_{\mathbb{P}}[\{M_t^f - M_s^f\} g(X_s)] = 0$$

meaning M_t^f is a martingale under $\mathbb{P}$ for arbitrary $f \in C_0^\infty(\mathbb{R}^d)$.

All that remains is to prove $\frac{A_h f}{h} \to \mathcal{L}f$ uniformly on compacts for every $f \in C_0^\infty(\mathbb{R}^d)$ and then using well posedness of MP we can conclude the result. Again let $f \in C_0^\infty(\mathbb{R}^d)$ and set

$$H(x, y) = \sum_{i=1}^d (y_i - x_i) \partial_i f(x) + \frac{1}{2} \sum_{i,j=1}^d (y_i - x_i)(y_j - x_j) \partial_{ij} f(x)$$

by Taylor

$$f(y) = f(x) + H(x, y) + R_3(y)$$

where
$$R_3(y) = \frac{1}{6}\sum_{i,j,k}(y_i - x_i)(y_j - x_j)(y_k - x_k)\partial_{ijk} f(\xi)$$
so
$$|f(y) - f(x) - H(x, y)| \le C|y - x|^3$$
Define $\mathcal{L}_h$ to be the second order differential operator with a^h and b^h the diffusion and drift coefficients respectively. Then

$$\begin{aligned}\mathcal{L}_h f(x) &= \frac{1}{2h}\sum_{i,j=1}^{d}\int_{|y-x|\le 1}(y_i - x_i)(y_j - x_j)\partial_{ij} f(x)\Pi_h(x, dy)\\ &+ \frac{1}{h}\sum_{i=1}^{d}\int_{|y-x|\le 1}(y_i - x_i)\partial_i f(x)\Pi_h(x, dy)\\ &= \frac{1}{h}\int_{|y-x|\le 1} H(x, y)\Pi_h(x, dy)\end{aligned}$$

therefore
$$\Big|\frac{1}{h}A_h f(x) - \mathcal{L}_h f(x)\Big| =$$
$$\Big|\frac{1}{h}\int\{f(y) - f(x)\}\Pi_h(x, dy) - \frac{1}{h}\int_{|y-x|\le 1} H(x, y)\Pi_h(x, dy)\Big| \le$$
$$\Big|\frac{1}{h}\int_{|y-x|\le 1}\{f(y)-f(x)-H(x, y)\}\Pi_h(x, dy)\Big|+\Big|\frac{1}{h}\int_{|y-x|>1}\{f(y)-f(x)\}\Pi_h(x, dy)\Big|$$
$$\le \frac{C}{h}\int_{|y-x|\le 1}|y - x|^3\Pi_h(x, dy) + \frac{1}{h}\int_{|y-x|>1}|f(y) - f(x)|\Pi_h(x, dy)$$
$$\le \frac{C}{h}\int_{\epsilon<|y-x|\le 1}|y - x|^3\Pi_h(x, dy) + \frac{C\epsilon^3}{h}\int_{|y-x|\le\epsilon}\Pi_h(x, dy) + 2||f||\Delta_h^1(x) \le$$
$$C\Delta_h^\epsilon(x) + C\epsilon^3(1 - \Delta_h^\epsilon(x)) + 2||f||\Delta_h^1(x)$$
for all $0 < \epsilon < 1$.
Because of condition (iii), we get
$$\Big|A_h f(x) - \mathcal{L}_h f(x)\Big| \overset{unif}{\to} 0$$
and by conditions (i) and (ii) we have
$$\mathcal{L}_h f \to \mathcal{L} f$$

uniformly on compacts.
Therefore

$$\frac{1}{h} A_h f \to \mathcal{L} f$$

uniformly on compacts.

We have that the MP for $\mathcal{L}$ is well posed. $\{\mathbb{P}_h^{x_0}\}_{h>0}$ is a relatively compact family of probability measures, that is that every sequence $\mathbb{P}^n$ of elements of $\{\mathbb{P}_h^{x_0}\}$ has a convergent subsequence. Every limit point of such subsequence is the unique solution $\mathbb{P}$ to the MP$(\mathcal{L}, \delta_{x_0})$. This means that every sequence of elements of $\{\mathbb{P}_h^{x_0}\}_h$ converges weakly to $\mathbb{P}$ and this means $\mathbb{P}_h^{x_0} \Rightarrow \mathbb{P}$.

References

[1] Edward. A. Azoff, Borel Measurability in Linear Algebra[1974], Proceedings of the American Mathematical Society, Volume 42, Number 2, February 1974.

[2] Richard F. Bass, Diffusions and Elliptic Operators[1997]. Springer New York, NY.

[3] Richard F. Bass, Probabilistic Techniques in Analysis[1995]. Springer New York, NY.

[4] Patrick Billingsley, Convergence of Probability Measures, Second Edition[1999], Wiley Series in Probability and Statistics.

[5] Nicolas Bourbaki, General Topology[1966], Springer-Verlag Berlin Heidelberg 1989.

[6] A. S. Cherny, Families of Consistent Probability Measures[2002], Theory of Probability Its Applications, Volume 46, Issue 1, January 2002.

[7] Stewart N. Ethier, Thomas G. Kurtz, Markov Processes: Characterization and Convergence[1986], Wiley Series in Probability and Statistics.

[8] Ioannis Karatzas, Steven E. Shreve, Brownian Motion and Stochastic Calculus[1987], Springer New York, NY.

[9] Daniel W. Stroock, S. R. Srinivasa Varadhan, Multidimensional Diffusion Processes[1979], Springer Berlin, Heidelberg.