This problem set exposes you to some ideas in stochastic processes with the aim of developing your intuition and understanding of SDEs as relevant to diffusion modelling. These problems are more mathematically advanced than anything you would be expected to do on the exam.

## Problem 1: From VP SDE to VE SDE through a variable change

Consider the following two SDE types, frequently used in the context of diffusion models:

$$\text{VP SDE}: \; dz_t = -\alpha_t z_t \, dt + \gamma_t \, d\omega_t, \qquad \text{VE SDE}: \; dy_t = \sigma_t \, d\omega_t.$$

In this problem, we will show that one can obtain VE SDE from VP SDE through a change of variables. To do this, we define a new process $y_t = z_t e^{-\int_0^t \alpha_s \, ds}$ and compute $dy_t$.

1. The Euler-Maruyama approximation to $z_{t+\Delta t}$ is

$$z_{t+\Delta t} = z_t - \alpha_t z_t \Delta t + \gamma_t \sqrt{\Delta t} \varepsilon + o(\Delta t), \quad \varepsilon \sim \mathcal{N}(0, I).$$

   Use the definition of $y_t$ to show that

$$y_{t+\Delta t} = (y_t - \alpha_t y_t \Delta t) e^{\int_t^{t+\Delta t} \alpha_s \, ds} + \gamma_t e^{\int_0^{t+\Delta t} \alpha_s \, ds} \sqrt{\Delta t} \varepsilon + o(\Delta t).$$

2. Use that $e^{\int_t^{t+\Delta t} \alpha_s \, ds} = 1 + \alpha_t \Delta t + O(\Delta t^2)$ to show that

$$y_{t+\Delta t} = y_t + \gamma_t e^{\int_0^{t+\Delta t} \alpha_s \, ds} \sqrt{\Delta t} \varepsilon + o(\Delta t).$$

3. The above equation is, up to $o(\Delta t)$, an Euler-Maruyama integration scheme for a VE SDE in $y_t$. What is this SDE's diffusion coefficient $\sigma_t$ in terms of $\alpha_t$ and $\gamma_t$?

## Problem 2: From Hierarchical VAE to SDE

Consider a noising scheme in which $z_n \sim \mathcal{N}(z_0, V_n)$ and each $z_{n+1}$ is obtained from $z_n$ by independent Gaussian increments. In this problem, we will show that this noising scheme defines the same distribution as the following SDE:

$$dz_t = \sqrt{\frac{\partial}{\partial t} V(t)} \, d\omega_t,$$

where $V : [0, N] \to \mathbb{R}$ is a monotonic differentiable function such that $V(n) = V_n$ for integers $n$.

1. Start by finding the closed form for the transition from $z_n$ to $z_{n+1}$.
2. Assume that we integrate the SDE from time $n$ to $n+1$ using the Euler-Maruyama scheme with step $\Delta t = \frac{1}{K}$. Show that the marginal distribution of $z_{n+1}$ given $z_n$ under this scheme is

$$z_{n+1} \sim \mathcal{N}\left(z_n, \frac{1}{K}\left(\frac{\partial V}{\partial t}(n) + \frac{\partial V}{\partial t}\left(n + \frac{1}{K}\right) + \cdots + \frac{\partial V}{\partial t}\left(n + \frac{K-1}{K}\right)\right)\right).$$

3. Take the limit $K \to \infty$ and show the variance in the Part 2 approaches $V(n+1) - V(n)$, recovering the expression in Part 1.

## Problem 3: From SDE to ODE flow

Consider the following forward SDE: $dz_t = \sigma_t \, d\omega_t$. Denote its marginal densities at time $t$ by $p_t$.

1. Use the identities from the lecture to write the corresponding reverse SDE and probability flow ODE.
2. Show that the Fokker-Planck-Kolmogorov (FPK) equation for the forward SDE, the FPK equation for the reverse SDE, and the continuity equation for the probability flow ODE all coincide.

## Problem 4: From ODE flow to Normalising Flows

In exercise 2 of Week 4 we showed that for a transition of the form $z_{t+\Delta t} = z_t + u_t(z_t)\Delta t$ – the form of an Euler integration step for an ODE – the change of density can be written as

$$\log p_{t+\Delta t}(z_{t+\Delta t}) - \log p_t(z_t) = -\mathrm{Tr}\left[\mathrm{J}_{u_t}\right]\Delta t + O(\Delta t^2),$$

where $J_{u_t}$ is the Jacobian of $u_t$ at $z_t$.
In this problem, we use this result to derive the continuity equation.

1. First, check using Taylor expansions that
$$\log p_{t+\Delta t}(z_{t+\Delta t}) = \log p_t(z_t) + \frac{\partial}{\partial t}\log p_t(z_t)\Delta t + \nabla \log p_t(z_t) \cdot (z_{t+\Delta t} - z_t) + O(\Delta t^2).$$
2. Using Part 1 and the transition formula, convert the change of density formula to
$$\frac{\partial}{\partial t}\log p_t(z_t) = -\nabla \cdot u_t - \nabla \log p_t(z_t) \cdot u_t(z_t).$$
You will need to use that $\mathrm{Tr}\left[\mathrm{J}_{u_t}\right] = \nabla \cdot u_t$.
3. Multiply both sides by $p_t$ and derive the usual form of the continuity equation.
(You will need the product rule for divergences: for a scalar function $f$ and vector field $v$, $\nabla \cdot (fv) = f(\nabla \cdot v) + (\nabla f) \cdot v$.)