EXERCISE DAY 1
CPDE SUMMER SCHOOL:
A PRACTICAL INTRODUCTION TO CONTROL, NUMERICS AND MACHINE LEARNING

DANIËL VELDMAN

Background

The following is a simplified version of the problem setting in

D.W.M. Veldman, R.H.B. Fey, H.J. Zwart, M.M.J. van de Wal, J.D.B.J. van den Boom, H. Nijmeijer (2021). Optimal thermal actuation for mitigation of heat-induced wafer deformation. *IEEE Transactions on Control System Technology.* 29(2), 514-529.

The wafer is a thin silicon disk, typically with a radius of 300 mm and a thickness of 0.7 mm. When the wafer is exposed to the projection light, it is placed on a water-cooled supporting structure which is assumed to have a constant temperature T_0. Because the wafer is thin, the temperature variations along the thickness are negligible and the temperature field in the wafer can be considered to be a function of the in-plane Cartesian coordinates (x, y) and time t only. The temperature increase in the wafer T w.r.t. the temperature of the supporting structure T_0 is the solution of the two-dimensional heat equation

$$\rho c H \frac{\partial T}{\partial t} = kH \left(\frac{\partial^2 T}{\partial x^2} + \frac{\partial^2 T}{\partial y^2} \right) - h_c T + Q, \tag{1}$$

where ρ [kg/m^3], c [J/kg/K], k [W/m/K], and H [m] are the mass density, specific heat capacity, thermal conductivity, and thickness of the wafer, respectively, h_c [W/m^2/K] is the thermal conductance between the wafer and the supporting structure, and $Q = Q(x, y, t)$ [W/m^2] is the heat load that results from the projecting light and from actuation, i.e.

$$Q = Q_{\text{exp}} + Q_{\text{act}}, \tag{2}$$

where $Q_{\text{exp}} = Q_{\text{exp}}(x, y, t)$ is the heat load resulting from the light that projects the pattern of electronic connections on the wafer and $Q_{\text{act}} = Q_{\text{act}}(x, y, t)$ is the actuation heat load. Note that convective and radiative heat transfer are negligible compared to the heat conduction to the supporting structure. It is assumed that the wafer temperature is initially equal to the temperature of the supporting structure T_0, i.e. the initial condition is $T(x, y, 0) = 0$. The spatial domain $(x, y) \in \mathbb{R}^2$ is considered to be infinite.

The heat load Q_{exp} is induced by the light that projects the pattern of electronic connections on the wafer and has a power P_{exp} [W] which is uniformly applied over the slit $\Omega_{\text{slit}} \subset \mathbb{R}^2$ (the red area in Fig. 1 with length L and width W). During the exposure of the wafer, the light source consecutively scans about 100 rectangular areas on the wafer, which are called fields. The scanning of a single

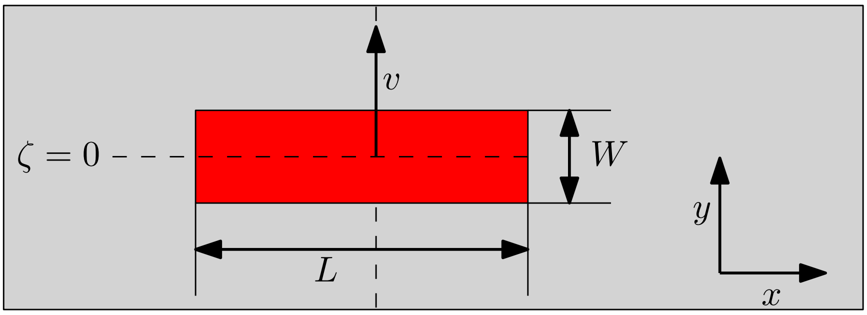

FIGURE 1. The heat load (red) that is applied to the wafer (gray)

field is considered. During the time interval $t \in (0, t_e)$ in which a single field is scanned, the slit moves with a constant velocity v in the positive y-direction,

$$Q_{\text{exp}}(x, y, t) = B_{\text{d}}(x, y - vt), \tag{3}$$

where $B_d(x, \zeta)$ is a multiple of the characteristic function of the set $[-L/2, L/2] \times [-W/2, W/2]$.

The actuation heat load has two spatial shapes that move together with $Q_{\text{exp}}(x, y, t)$,

$$Q_{\text{act}}(x, y, t) = B_1(x, y - vt)u_1(t) + B_2(x, y - vt)u_2(t), \tag{4}$$

where $B_1(x, \zeta)$ is a multiple of the characteristic function of the set $[-L/2, L/2] \times [W/2, 3W/2]$, $B_2(x, \zeta)$ is a multiple of the characteristic function of the set $[-L/4, L/4] \times [-3W/2, -W/2]$, and $u_1(t)$ and $u_2(t)$ are the controls.

Because the considered domain is infinite and the applied heat load is moving, it is convenient to consider a moving coordinate system $(x, \zeta, t) = (x, y - vt, t)$ in which the shape of Q_{exp} is fixed. Let $T^{(y)}(x, y, t)$ and $Q^{(y)}(x, y, t)$ denote the temperature field and applied heat load expressed in (x, y, t)-coordinates as in (1). The temperature field $T^{(\zeta)}(x, \zeta, t)$ and applied heat load $Q^{(\zeta)}(x, \zeta, t)$ expressed in (x, ζ, t)-coordinates are then equal to $T^{(y)}(x, \zeta + vt, t)$ and $Q^{(y)}(x, \zeta + vt, t)$, respectively. It can be shown that $T(x, \zeta, t) = T^{(\zeta)}(x, \zeta, t)$ satisfies

$$\rho c H \left(\frac{\partial T}{\partial t} - v \frac{\partial T}{\partial \zeta} \right) = kH \left(\frac{\partial^2 T}{\partial x^2} + \frac{\partial^2 T}{\partial \zeta^2} \right) - h_c T + B_{\text{d}} + B_1 u_1 + B_2 u_2. \tag{5}$$

In the remainder, only the (x, ζ, t)-coordinate system will be used and the used coordinate system will no longer be indicated, $T = T(x, \zeta, t) = T^{(\zeta)}(x, \zeta, t)$ and $Q = Q(x, \zeta, t) = Q^{(\zeta)}(x, \zeta, t)$. Note that the origin of (x, ζ)-coordinate system is at the center of the slit (see Fig. 1).

Our goal is to compute the controls $u_1(t)$ and $u_2(t)$ that minimize

$$J = 10^8 \int_0^{t_e} \iint_{\Omega_{\text{slit}}} (T(x, \zeta, t))^2 \, \mathrm{d}x \, \mathrm{d}\zeta \, \mathrm{d}t + \int_0^{t_e} \left(u_1^2(t) + u_2^2(t) \right) \, \mathrm{d}t \tag{6}$$

EXERCISE

A finite element discretization (on a truncated spatial domain) yields an optimal control problem of the form where

$$J = \frac{1}{2} \int_0^T (\mathbf{x}(t))^\top \mathbf{Q} \mathbf{x}(t) + \mathbf{u}(t)^\top \mathbf{R} \mathbf{u}(t) \, \mathrm{d}t, \tag{7}$$

should be minimized subject to the dynamics

$$\mathbf{E}\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{B}_{\text{d}} + \mathbf{B}\mathbf{u}(t), \qquad \mathbf{x}(0) = \mathbf{x}_{\text{init}}.$$

The matrices $\mathbf{E}$, $\mathbf{A}$, $\mathbf{B}$, $\mathbf{x}_{\text{init}}$, $\mathbf{Q}$, and $\mathbf{R}$ are computed in the first 82 lines of `CPDESS_Exercise1` and $\mathbf{u}(t) = [u_1(t), u_2(t)]^\top$. You do not need to study these lines in detail.

Note: if you are using Octave instead of Matlab you should use the files ending on `_octave` when these are available. Files for which there is no duplicate ending on `_octave` should work both in Matlab and Octave.

a. Compute the state $\mathbf{x}(t)$ resulting from an input $\mathbf{u}(t)$ using the Crank-Nicolson scheme by filling in the missing lines in `OCP_compute_temperature`. Use a time grid with $N_T = 151$ points. You can precompute an LU factorization to speed up computations. You can visualize the obtained solution by uncommenting line 99 in `CPDESS_Exercise1`.
b. Use the obtained solution to evaluate the cost functional J according to the scheme explained in the lecture by completing the missing line in `OCP_costfunction`. Also compute to the gradient of J based on the adjoint state $\boldsymbol{\varphi}(t)$ by completing the missing lines in `OCP_compute_adjoint`. Follow the procedure discretize-then-optimize approach given in the lecture.
c. Compute the coefficients G and H in the quadratic approximation of the cost functional $\beta \mapsto J(u_0 - \beta \nabla J)$ by completing the files `OCP_innerproduct`, `OCP_compute_dtemperature`, and `OCP_hessian`. Compare the obtained quadratic approximation to the true values of the cost functional. Do you expect to see any difference?
d. Use the results from parts a, b, and c to develop and implement a gradient-based algorithm to minimize the functional J by completing the missing lines in `CPDESS_Exercise1`.
   Plot the resulting optimal controls $u_1^*(t)$ and $u_2^*(t)$.