# Optimal control in octopus: status report

Alberto Castro

ARAID Foundation and Institute for Biocomputation and Physics of Complex Systems (BIFI), Zaragoza (Spain)

octopus meeting, Jena, 26-30.01.2015



Instituto Universitario de Investigación
Biocomputación y Física
de Sistemas Complejos
Universidad Zaragoza

# Outline

# Outline

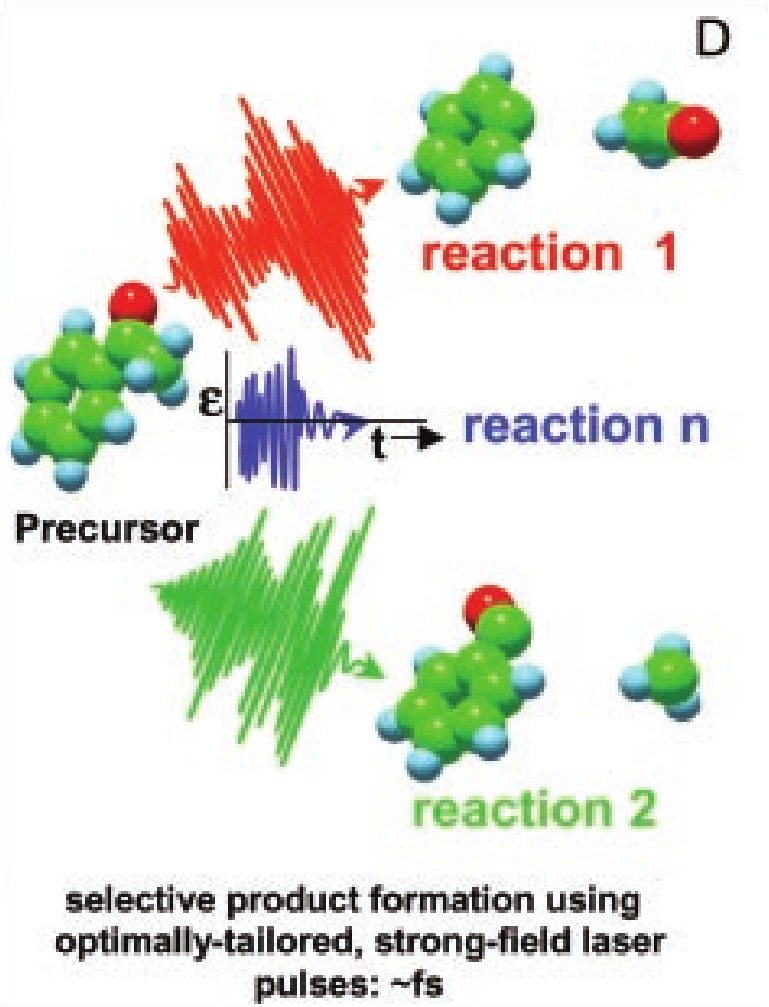
# "Standard" ref

**Control of quantum phenomena: past, present and future**

**Constantin Brif, Raj Chakrabarti[1] and Herschel Rabitz[2]**
Department of Chemistry, Princeton University, Princeton, NJ 08544, USA
E-mail: cbrif@princeton.edu, rchakra@purdue.edu and hrabitz@princeton.edu

*New Journal of Physics* **12** (2010) 075008 (68pp)

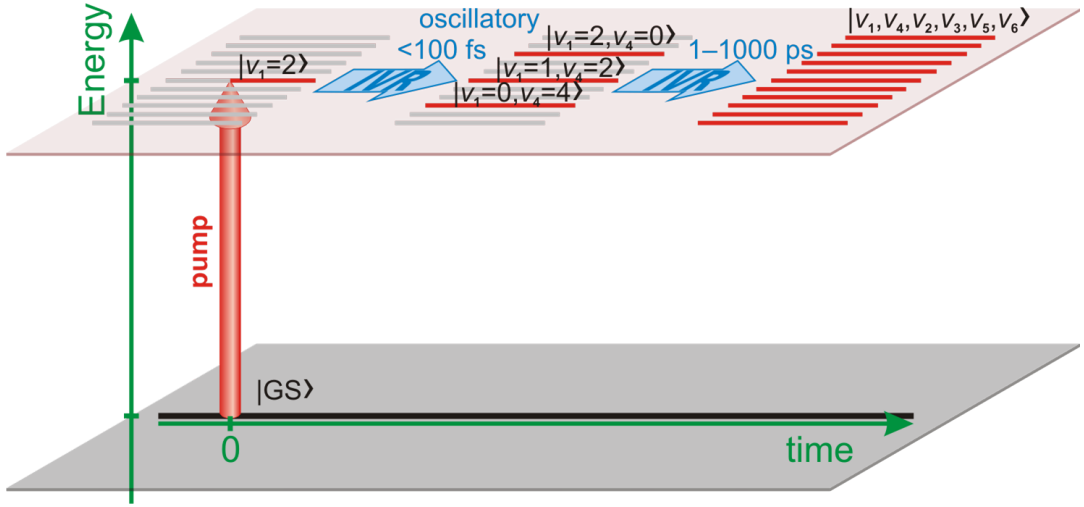
# Basic idea, in a cartoon




[R. J. Levis *et al*, Science **292**, 709 (2001)]

# Mono-chromatic lasers: IVR

- Lasers (coherent, monochromatic, intense light) promised to deliver precise control of quantum systems
- Initially, the first attempts to control molecules (i.e. "photo-selective chemistry") were based on tuning the laser frequency to specific bonds
- Those attempts were seldom successful, due to "intramolecular vibrational redistribution".

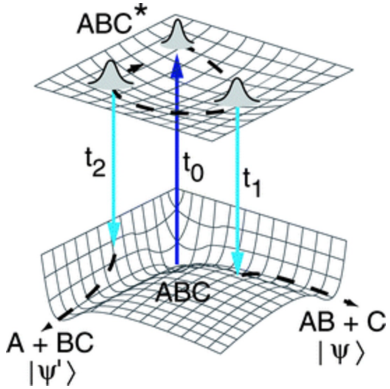



- Analogous problems will appear in other quantum control attempts, beyond molecular photo-chemistry.

# Pioneering schemes

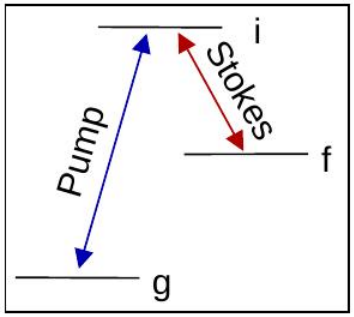
- **Interferences, and the "two pathway" scheme** Use of two monochromatic lasers with commensurate frequencies for creating quantum interference between two reaction pathways [P. Brumer and M. Shapiro, Chem. Phys. Lett. **126**, 541 (1986)].

- **Pump and dump** [D. J. Tannor and S. A. Rice J. Chem. Phys. **83**, 5013 (1985)]




Limitations: knowledge of the potential energy surfaces, competing processes.

- **STIRAP: STImulated Raman scattering involving Adiabatic Passage** [K. Bergmann *et al* Chem. Phys. Lett. **149**, 463 (1988)]




STIRAP

# Adaptive feedback control

VOLUME 68, NUMBER 10 PHYSICAL REVIEW LETTERS 9 MARCH 1992

## Teaching Lasers to Control Molecules

Richard S. Judson[a]

*Center for Computational Engineering, Sandia National Laboratories, Livermore, California 94551-0969*

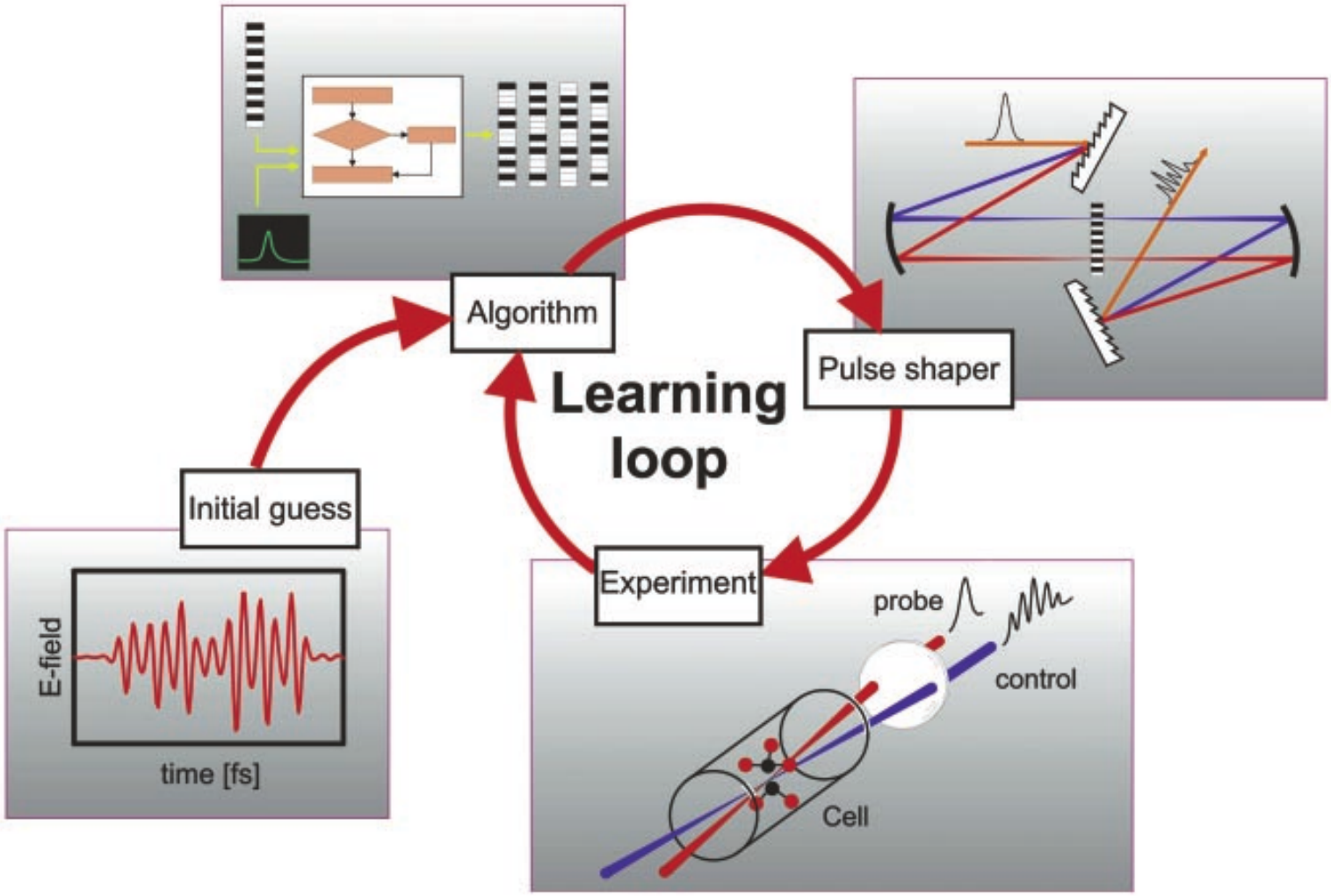
Herschel Rabitz

*Department of Chemistry, Princeton University, Princeton, New Jersey 08544*

(Received 26 August 1991)

We simulate a method to teach a laser pulse sequences to excite specified molecular states. We use a learning procedure to direct the production of pulses based on "fitness" information provided by a laboratory measurement device. Over a series of pulses the algorithm learns an optimal sequence. The experimental apparatus, which consists of a laser, a sample of molecules, and a measurement device, acts as an analog computer that solves Schrödinger's equation *exactly*, in real time. We simulate an apparatus that learns to excite specified rotational states in a diatomic molecule.

# The learning loop




[H. Rabitz *et al*, Science **288**, 824 (2000)]

# Examples of AFC experiments: other

- Photo-dissociation reactions in molecules.
- High harmonic generation.
- Multi-photon ionization of atoms.
- Electronic excitation in molecules (fluorescence is used as the probe to build the merit function).
- Molecular alignment.
- Photo-induced electron transfer between molecules
- Photo-isomerization of molecules.
- etc.

# What optimal control theory is, in a nutshell.

Typical formulation of a (general) optimal control problem:

- Dynamical system:

$$\begin{aligned} \dot{x}(t) &= f(x(t), u(t), t) \\ x(0) &= x_0 \end{aligned}$$

  Typically, $u = u(t)$. But it can be a set of parameters whatsoever.

- Minimize the cost functional:

$$F[x, u] = F^{\text{terminal}}[x(T), u] + \int_0^T \mathrm{d}t \; L(x(t), u(t)]$$

- Since $u \to x[u]$, it amounts to minimizing

$$G[u] = F[x[u], u]$$

# What optimal control theory is, in a nutshell.

Problems to be answered by OCT:

- Design a good algorithm to find the optimal $u$.
  For example, find an efficient way to compute $\nabla G[u]$.
- How many solutions are there? Are there local minima?
- How stable are the minima with respect to small changes in $u$? (find $\nabla^2 G[u]$)
- How stable are the minima with respect to small changes in the problem definition (changes in $f$ or $x_0$)?
- Is the problem *controllable* at all?

# Essential theoretical results

- Pontryagin's minimum principle (1956)
  [V.G. Boltyanskii, R.V. Gamkrelidze, and L.S. Pontryagin, "Towards a theory of optimal processes", (Russian), Reports Acad. Sci. USSR **110**, 1 (1956)]
  It provides a *necessary* condition for the minimum – in practice, typically, an expression for $\nabla G[u]$ so that the equation $\nabla G[u] = 0$ can be posed.
- Hamilton-Jacobi-Bellman equation (1954)
  (Theory of "dynamic programming", Richard Bellman)
  [R.E Bellman, "Dynamic Programming and a new formalism in the calculus of variations" Proc. Nat. Acad. Sci. **40**, 231 (1954)]

# Hamiltonian systems, calculation of the gradient

$$\begin{aligned} \dot{q}_i(t) &= \frac{\partial H}{\partial p_i}(q(t), p(t), u, t)\,, \\ \dot{p}_i(t) &= -\frac{\partial H}{\partial q_i}(q(t), p(t), u, t)\,. \end{aligned}$$

In condensed vector notation:

$$\begin{aligned} \dot{\mathbf{q}} &= \frac{\partial H}{\partial \mathbf{p}}\,, \\ \dot{\mathbf{p}} &= -\frac{\partial H}{\partial \mathbf{q}}\,. \end{aligned}$$

In this case, $x$ is the vector $[\mathbf{q}, \mathbf{p}]^\dagger$.
The cost (or target) functional can be, for example, any expression in the form:

$$F[\mathbf{q}, \mathbf{p}, u] = J_1[\mathbf{q}(T), \mathbf{p}(T)] + J_2[u]\,.$$

# Hamiltonian systems, calculation of the gradient

$$\frac{\partial G}{\partial u_m} = \frac{\partial J_2}{\partial u_m} - \int_0^T \mathrm{d}t\, \tilde{\mathbf{p}}(t) \cdot \frac{\partial}{\partial u_m} \frac{\partial H}{\partial \mathbf{p}}(\mathbf{q}(t), \mathbf{p}(t), u, t) + \int_0^T \mathrm{d}t\, \tilde{\mathbf{q}}(t) \cdot \frac{\partial}{\partial u_m} \frac{\partial H}{\partial \mathbf{q}}(\mathbf{q}(t), \mathbf{p}(t), u, t)$$

The "costate" $[\tilde{\mathbf{q}}, \tilde{\mathbf{p}}]^\dagger$ is itself a Hamiltonian system, determined by the quadratic Hamiltonian:

$$\tilde{H}(\tilde{q}, \tilde{p}, q, p, u, t) = \frac{1}{2}\tilde{\mathbf{q}}^t \mathbf{H}^{qq}(q, p, u, t)\tilde{\mathbf{q}} + \frac{1}{2}\tilde{\mathbf{q}}^t \mathbf{H}^{qp}(q, p, u, t)\tilde{\mathbf{p}} + \frac{1}{2}\tilde{\mathbf{q}}^t \mathbf{H}^{pq}(q, p, u, t)\tilde{\mathbf{p}} + \frac{1}{2}\tilde{\mathbf{q}}^t \mathbf{H}^{pp}(q, p, u, t)\tilde{\mathbf{p}}\,.$$

# Hamiltonian systems, calculation of the gradient

The matrices $\mathbf{H}^{xy}$ are defined as:

$$
\begin{aligned}
H_{ij}^{qq} &= \frac{\partial^2 H}{\partial q_i \partial q_j}(q,p,u,t) \\
H_{ij}^{qp} &= \frac{\partial^2 H}{\partial q_i \partial p_j}(q,p,u,t) \\
H_{ij}^{pq} &= \frac{\partial^2 H}{\partial p_i \partial q_j}(q,p,u,t) \\
H_{ij}^{pp} &= \frac{\partial^2 H}{\partial p_i \partial p_j}(q,p,u,t)
\end{aligned}
$$

# Quantum problems, calculation of the gradient

$$\hat{H} = \hat{H}[u_1, \ldots, u_M; t]$$

$$\begin{aligned} i\frac{\mathrm{d}}{\mathrm{d}t}|\Psi(t)\rangle &= \hat{H}[u;t]|\Psi(t)\rangle \\ |\Psi(t_0)\rangle &= |\Psi_0\rangle \end{aligned}$$

$$\Psi(t_0) \longrightarrow \Psi[u](t) \longrightarrow \Psi[u](T)$$

Maximize a quantity

$$F = F[\Psi[u](t)]\,,$$

that depends on the system evolution, or final state, or both.

# Quantum problems, calculation of the gradient

$$F[\Psi, u] = J_1[\Psi(T)] + J_2[u]$$

$$G[u] = F[\Psi[u], u]$$

$$\frac{\partial G}{\partial u_m} = \frac{\partial J_2}{\partial u_m} + 2\mathrm{Im}\int_0^T \mathrm{d}t \langle \chi(t)| \frac{\partial \hat{H}}{\partial u_m} |\Psi(t)\rangle \,,$$

where the "costate" $\chi$ verifies:

$$i\frac{\mathrm{d}}{\mathrm{d}t}|\chi(t)\rangle = \hat{H}(t)|\chi(t)\rangle \,,$$

$$|\chi(T)\rangle = \frac{\delta}{\delta \Psi^*(T)} F[\Psi(T)]$$

# Outline

# Outline

# Different dynamical systems: (1) "Standard" Schrödinger equation.

Example: [Zhu *et al.*, J. Chem. Phys. **108**, 1953 (1999)] (`06-zbr98.test`)
Transition between the ground and one excited vibrational state in a Morse potential:

$$V(r) = D_0[e^{-\beta(r-r_0)}-1]^2-D_0$$

$$H(t) = T_0+V(r)-\mu_0 r e^{-r/r^*}\epsilon(t$$

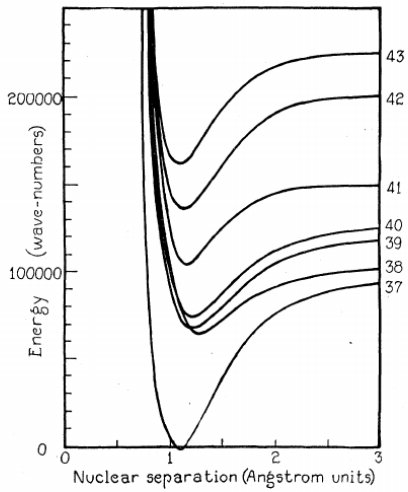



Fig. 1. Potential energy curves for nitrogen. Energy in wave-number and nuclear separation in Angstrom units. Numbers on curves refer to Table I.

[Morse, Phys. Rev. B **34**, 57 (1929)]

# Different dynamical systems: (1) "Standard" Schrödinger equation.

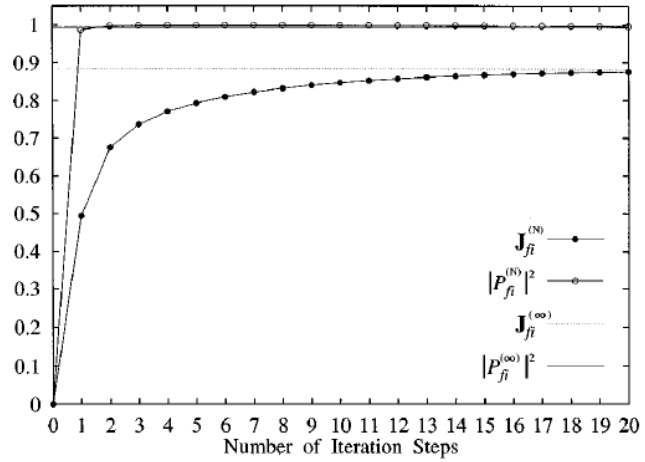



FIG. 1. Optimized objective functional $\mathbf{J}_{fi}^{(N)}$ and transition probability $|P_{fi}^{(N)}|^2$ for the transition of $v=0 \rightarrow v=1$. Rapid monotonic convergence is found.

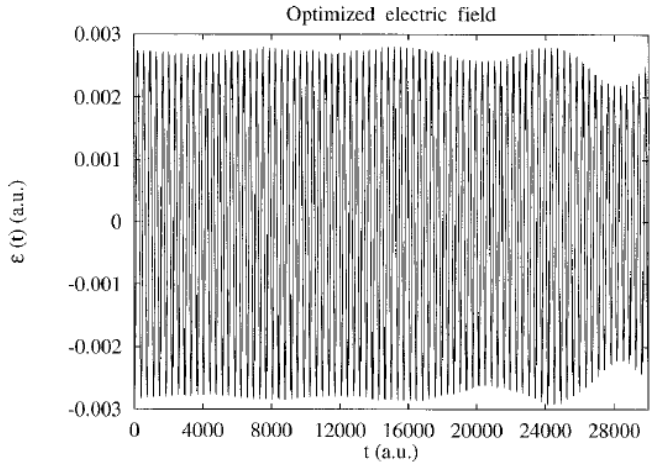



FIG. 2. Optimized electric field time dependence for the transition of $v=0 \rightarrow v=1$. It is almost a pure cosine function.

# Different dynamical systems: (2) TDDFT.

- We have a system of $N$ electrons, driven by an external potential $v_{\text{ext}}(\vec{r}, t, u)$.
- The time-dependent density is therefore determined by $u$:

$$u \longrightarrow n[u](\vec{r}, t) = \langle \Psi[u](t) | \hat{n}(\vec{r}) | \Psi[u](t) \rangle$$

- The objective is to maximize some function $G$ of the *control parameters* $u$, defined in terms of a functional of the density:

$$G[u] = \tilde{F}[n[u], u] \, .$$

- Since the definition is given in terms of the density, everything can be reformulated for the Kohn-Sham system, and the optimization will be equivalent. Since we use the Kohn-Sham substitution, we may use the Kohn-Sham orbitals instead:

$$F[\underline{\varphi}[u], u] \equiv \tilde{F}[n[u], u] \, , \quad n[u](\vec{r}, t) = \sum |\varphi_i[u](\vec{r}, t)|^2 \, .$$

# Different dynamical systems: (2) TDDFT.

- Definition of a target in terms of the Kohn-Sham orbitals:

$$F = F[\underline{\varphi}, u] \quad \Rightarrow \quad G[u] = F[\underline{\varphi}[u], u] = \tilde{F}[n[u], u] \, .$$

- Optimal control theory equations for TDDFT (terminal target only):

$$\begin{aligned} \nabla_u G[u] \quad = \quad & \nabla_u F[\underline{\varphi}[u], u] + \\ & 2\mathrm{Im} \left[ \sum_{i=1}^{N} \int_0^T \mathrm{d}t \, \langle \lambda_i[u](t) | \nabla_u \hat{H}[n[u](t), u, t] | \varphi_i[u](t) \rangle \right] \end{aligned}$$

$$\begin{aligned} \dot{\underline{\varphi}}[u](t) &= -i \underline{\underline{\hat{H}}}[n(t), u, t] \underline{\varphi}[u](t) \, , \\ \underline{\varphi}_u(0) &= \underline{\varphi}_0 \, , \\ \dot{\underline{\lambda}}[u](t) &= -i \left[ \underline{\underline{\hat{H}}}[n(t), u, t] + \underline{\underline{\hat{K}}}[\underline{\varphi}[u](t)] \right] \underline{\lambda}[u](t) \, , \\ \underline{\lambda}[u](T) &= \frac{\delta F}{\delta \underline{\varphi}^*}[\underline{\varphi}[u](T), u] \, . \end{aligned}$$

# Different dynamical systems: (2) TDDFT.

$$\dot{\underline{\lambda}}[u](t) = -i \left[ \underline{\underline{\hat{H}}}^{\dagger}[n[u](t), u, t] + \underline{\underline{\hat{K}}}[\underline{\varphi}[u](t)] \right] \underline{\lambda}[u](t) \,,$$

$$\dot{\lambda}_i[u](t) = -i\hat{H}^{\dagger}[n[u](t), u, t]\lambda_i[u](t) - i\sum_{j=1}^{N} \hat{K}_{ij}[\underline{\varphi}[u](t)]\lambda_j[u](t)$$

$$\langle \vec{r}|\hat{K}_{ij}[\underline{\varphi}[u](t)]|\lambda_j[u](t)\rangle =$$

$$-2i\varphi_i[u](\vec{r},t)\mathrm{Im}\left[\int \mathrm{d}^3r' \lambda_j[u]^*(\vec{r}',t) f_{\mathrm{Hxc}}[n[u](t)](\vec{r},\vec{r}')\varphi_j[u](\vec{r}',t)\right]$$

$$f_{\mathrm{Hxc}}[n[u](t)](\vec{r},\vec{r}') = \frac{1}{|\vec{r}-\vec{r}'|} + f_{\mathrm{xc}}[n[u](t)](\vec{r},\vec{r}')$$

# Different dynamical systems: (3) Ehrenfest dynamics, with TDDFT (or not).

Coupled electron-ion model: Ehrenfest dynamics:

$$\hat{H}[q,p,u,t] = H_{\rm clas}[q,p,u,t]\hat{I} + \hat{H}_{\rm quantum}[q,p,u,t]\,.$$

$$\begin{aligned}
\dot{q}_a(t) &= \frac{\partial H_{\rm clas}}{\partial p_a}[q(t),p(t),u,t] \\
&\quad + \langle\Psi(t)|\frac{\partial \hat{H}_{\rm quantum}}{\partial p_a}[q(t),p(t),u,t]|\Psi(t)\rangle \\
\dot{p}_a(t) &= -\frac{\partial \hat{H}_{\rm clas}}{\partial q_a}[q(t),p(t),u,t] \\
&\quad - \langle\Psi(t)|\frac{\partial \hat{H}_{\rm quantum}}{\partial q_a}[q(t),p(t),u,t]|\Psi(t)\rangle \\
\dot{\Psi}(x,t) &= -{\rm i}\hat{H}_{\rm quantum}[q(t),p(t),u,t]\Psi(x,t)\,,
\end{aligned}$$

[Castro and Gross, J. Phys. A **47**, 025204 (2014)]

# Outline

# Different algorithms: (1) Gradient-less algorithms

- Simpler approaches: *direct* or *gradient-less* algorithms. They only require a means to compute $G[u]$ (i.e. a method to propagate the dynamical equation and compute the resulting cost or target functional).
  - The most fashionable, the families of *evolutionary* or *genetic* algorithms.
  - Our choices:
    - The *simplex* algorithm [J.A. Nelder and R. Mead, Computer Journal **7**, 308 (1965)], and
    - the NEWUOA algorithm [M. J. D. Powell, IMA J. Numer. Anal. **28**, 649 (2008)].

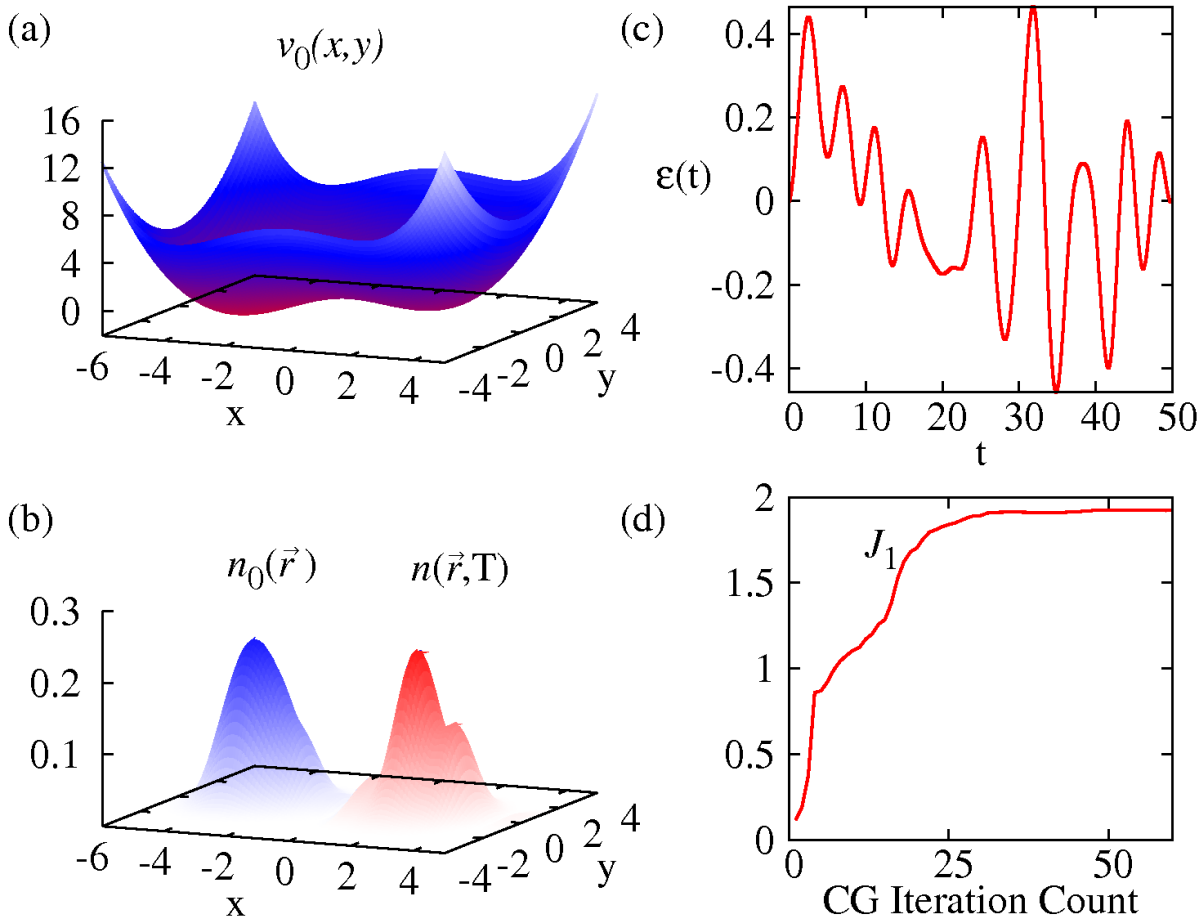
# Different algorithms: (2) "QOCT-specific" algorithms designed for real-space parametrizations.

- ZR98: [Zhu and Rabitz, J. Chem. Phys. **109**, 385 (1998)]
- ZBR98: [Zhu, Botina, and Rabitz, J. Chem. Phys. **108**, 1953 (1998)]
- Krotov: [V. F. Krotov, Automat. Remote Control **34**, 1863 (1973)]
- MT03: [Maday and Turinici, J. Chem. Phys. **118**, 8191 (2003)]
- WG05: [Werschnik and Gross, J. Opt. B **7**, S300 (2005)]

All of these assume that the parameters $u_1, \ldots, u_M$ are the real-time discretization values of the control function $\varepsilon(t)$: $\varepsilon(t_1), \varepsilon(t_2), \ldots, \varepsilon(t_M)$. It is difficult to establish constraints.

# Different algorithms: (3) The usual BFGS

In fact, one can use all the gradient-based methods implemented in the GSL library (conjugate gradients, etc.). The parametrization of the control field can then be arbitrary, although only some options have been implemented.




[Castro and Gross, PRL **109**, 153603 (2012)]

# Outline

# Different targets

- State-to-state transitions.

$$J_1[\Psi] = \langle \Psi(T) | \Psi_O \rangle \langle \Psi_O | \Psi(T) \rangle \, .$$

- "Ground-state transformation"

$$\Psi = \det[\psi_1, \ldots \psi_N]$$

$$|\Psi_O\rangle = \hat{a}^\dagger_{a_1} \hat{a}_{i_1} \ldots \hat{a}^\dagger_{a_k} \hat{a}_{i_k} |\Psi\rangle$$

- "Excited state" population

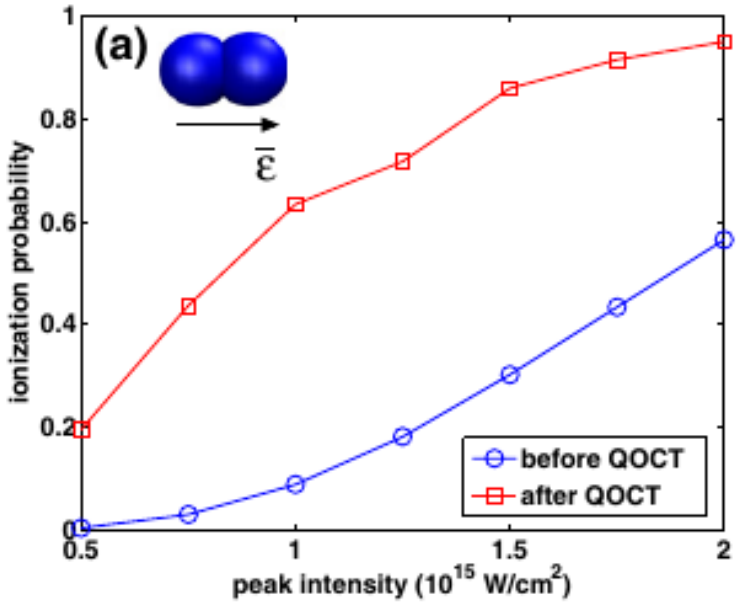
$$|\Psi_O\rangle = \sum_{ai} c_{ai} \hat{a}^\dagger_a \hat{a}_i |\Psi\rangle$$

# Different targets

- Ionization

$$J_1[\Psi] = 1 - \sum_i |\langle \Psi(T) | \Psi_i \rangle|^2 .$$

$$J_1[\Psi] = \int_{r \in U} \mathrm{d}^3_r n(\vec{r}, T) .$$




[Castro *et al.*, EPL **87**, 53001 (2009)]

# Different targets

- Density, current

$$J_1[\Psi] = \alpha \int \mathrm{d}^3 r \; a(n(\vec{r}, T)) + \beta \int \mathrm{d}^3 r \; b(\vec{j}(\vec{r}, T)) \, .$$

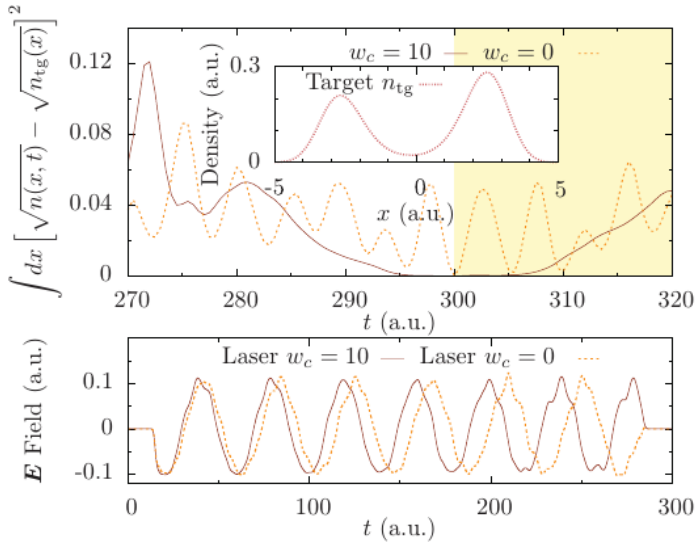



FIG. 1. (Color online) 1D QW. The overlap between the controlled density $n$ and the target density $n_{\rm tg}$ is shown in the top panel. For minimized current ($w_c = 10$) the overlap with the target (found in the inset) at terminal time $T = 300$ a.u. is better. Additionally, density $n$ is significantly more stable than without current control ($w_c = 0$). The respective laser pulses are reported in the lower panel.

[Kammerlander *et al.*, PRA **83**, 043413 (2011)]

# Different targets

- Time-dependent targets.

$$J_1[\Psi] = \int_0^T \mathrm{d}t\, G(\Psi(t))\,.$$

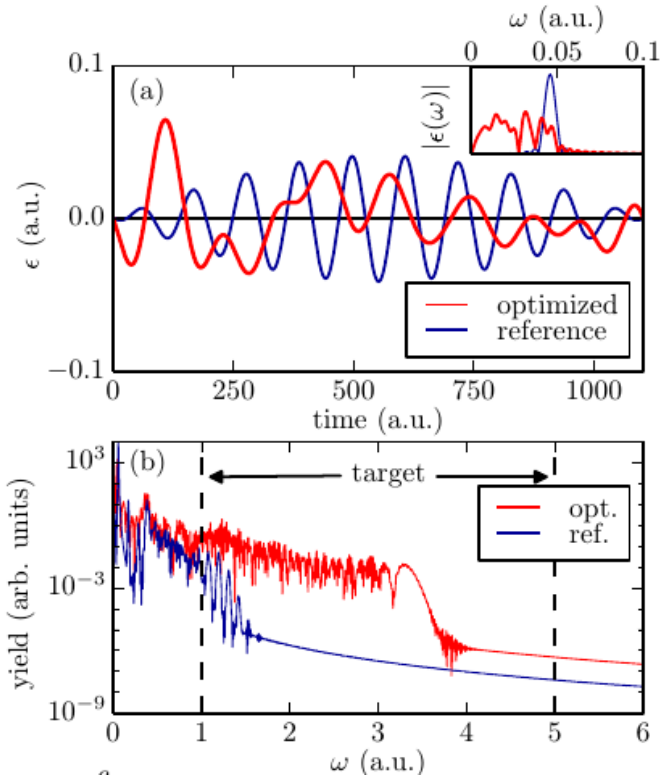
In these cases, the backwards propagation includes an inhomogeneous term.

# Different targets

- High-harmonic generation.

$$J_1[\Psi] = \int_0^\infty \mathrm{d}\omega\ \alpha(\omega) H(\omega)\,,$$

$$H(\omega) = \int_0^T \mathrm{d}t\ |e^{-i\omega t} \frac{\mathrm{d}^2}{\mathrm{d}t^2} x(t)|^2\,.$$




[Solanpää *et al.*, PRA **90**, 053402 (2014)]

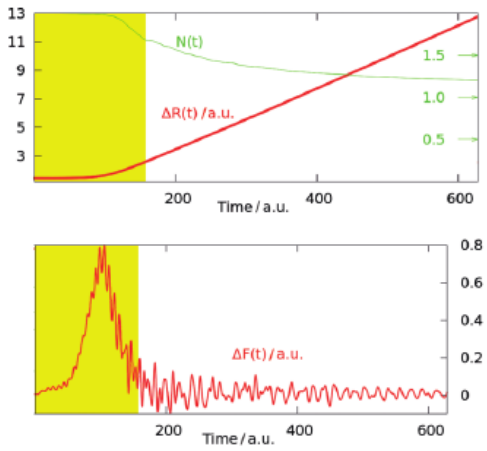
# Different targets

- Quantum-classical targets.

$$F = F[q, p, \Psi]$$

For example, for the dissociation of two atoms:

$$F = F[\vec{R}_1, \vec{R}_2, \vec{P}_1, \vec{P}_2, \Psi] = (\vec{R}_1 - \vec{R}_2)^2 \, .$$




**Ultrafast and optimal:** optimal ultrashort pulses are theoretically constructed by the combination of time-dependent density functional theory and optimal control, optimizing the classical opposing force on the nuclei for the hydrogen molecule (see picture; top panel: internuclear distance and the bound electrons count; bottom panel: opposing force; gray region: pulse duration).

[Castro, ChemPhysChem **14**, 1488 (2014)]

# Outline

# Different parametrizations

- Real-time parametrization

$$u_1, u_2, \cdots = \varepsilon(t_1), \varepsilon(t_2), \ldots$$

- Fourier coefficients

$$u_1, u_2, \cdots = \{a_n, b_n\}\,.$$

$$\varepsilon(t) = a_0 + \sum_n a_1 \cos(w_n t) + b_n \sin(w_n t)$$

One may want to force:

$$\int_0^T \mathrm{d}t\; \varepsilon(t) = 0 \Longrightarrow a_0 = 0\,,$$

$$\varepsilon(0) = \varepsilon(T) = 0 \Longrightarrow \sum_n a_n = 0\,.$$

- Constant-fluence constraint.

$$\int_0^T \mathrm{d}t\; \varepsilon^2(t) = C\,,$$

In this case, $u_1, u_2, \ldots$ are certain hyper-spherical angles that are obtained after some transformations.

# Outline

# Summary

| Systems | Algorithms | Targets | Parametrizations |
|---|---|---|---|
| Schrödinger | ZR98 | State projection | Real-time |
| TDKS | ZBR98 | Density | Fourier |
| Ehrenfest + wave function | WG05 | Current | Fourier (constrained) |
| Ehrenfest + TDDFT | MT03 | HHG | |
| | Krotov | Ionization | |
| | BFGS, CG (GSL) | Classical vars. | |
| | Nelder-Mead | Spin | |
| | NEWUOA | | |