

finalmente en la sección 7 se formulan las conclusiones del trabajo.

2. PRELIMINARES

2.1 Definiciones

Un problema de optimización multi-objetivo consiste en minimizar de manera simultánea un conjunto de funciones objetivo de la forma

$$\begin{aligned} \min_{\mathbf{x} \in \mathcal{X}} f_1(\mathbf{x}) \\ \min_{\mathbf{x} \in \mathcal{X}} f_2(\mathbf{x}) \\ \vdots \\ \min_{\mathbf{x} \in \mathcal{X}} f_M(\mathbf{x}) \end{aligned} \quad (1)$$

donde \mathbf{x} representa el vector de variables de decisión y \mathcal{X} el espacio de búsqueda o espacio de variables.

En general, la reducción del valor de una función objetivo f_i puede conducir al incremento en otra(s), por lo cual no existe un $\mathbf{x}^* \in \mathcal{X}$ que sea mínimo de todas simultáneamente. Para Osyczka (1985) el término "min" debe en este caso ser interpretado como la búsqueda de soluciones aceptables (a criterio del diseñador) para todos los objetivos. Es decir, resolver el problema (1) equivale a **determinar los mejores compromisos posibles**.

La metodología tradicional para resolver este tipo de problemas se basa en utilizar una transformación escalar, con la esperanza de que la solución obtenida (mediante cualquier método de optimización escalar disponible) satisfaga las especificaciones de diseño. Kallianmoy Deb (2001) se refiere a este tipo de enfoques como "*artificial fix-ups*", debido a que en general sólo producen una solución, siendo necesaria repetir la búsqueda varias veces para identificar los buenos compromisos. Por otra parte, necesitan conocimiento previo del problema (para definir las prioridades entre los objetivos, por ejemplo).

El otro enfoque que se ha sugerido para resolver el problema multi-objetivo consiste en obtener un subconjunto de soluciones caracterizadas por la propiedad de que sea imposible reducir el valor de cualquier objetivo sin incrementar al menos otro. Cuando dicho conjunto existe, el mismo se conoce como *Conjunto Óptimo de Pareto*, tal como se define a continuación.

Sea el problema de optimización planteado en (1). Se denomina **Conjunto Óptimo Global de Pareto (COGP)**, y se denota \mathcal{P}^* , al conjunto de $\mathbf{x}^* \in \mathcal{X}$ tales que $\nexists \mathbf{x} \in \mathcal{X}$ que cumpla

$$f_i(\mathbf{x}) \leq f_i(\mathbf{x}^*), \forall i \in \{1, \dots, M\} \quad (2)$$

$$\exists j \in \{1, \dots, M\} : f_j(\mathbf{x}) < f_j(\mathbf{x}^*) \quad (3)$$

Se denomina **espacio objetivo** y se denota \mathcal{F} al conjunto de vectores \mathbf{z} tales que

$$\exists \mathbf{x} \in \mathcal{X} : \mathbf{z} = (f_1(\mathbf{x}) \ f_2(\mathbf{x}) \ \dots \ f_M(\mathbf{x}))^T \quad (4)$$

Sean z^1 y z^2 dos vectores pertenecientes a \mathcal{F} . Se dice entonces que z^2 es **dominado** por z^1 y se denota $z^1 \preceq z^2$ si y solo si

$$\begin{aligned} z_i^1 \leq z_i^2, \quad i = 1, 2, \dots, M \\ \exists i \in \{1, \dots, M\} : z_i^1 < z_i^2 \end{aligned} \quad (5)$$

En lo que sigue se denota \mathcal{P} a una aproximación cualquiera del **COGP** en el espacio de variables y \mathcal{PF} a la aproximación correspondiente en el espacio objetivo. Para una explicación detallada y actualizada acerca del tema es posible consultar la monografía publicada por Coello et al. (2007).

2.2 Indicadores de Performance

El problema de la evaluación de distintas aproximaciones del **COGP**, es un tema que se encuentra actualmente en desarrollo. Existe bastante consenso con respecto a tres aspectos que se deben evaluar:

- Cantidad de soluciones no-dominadas.
- Distancia entre las soluciones de la aproximación y las soluciones del **COGP**.
- Distribución de las soluciones: se desea obtener una distribución tan uniforme como sea posible.

A continuación se definen los indicadores de performance que se utilizaron en este trabajo. Es interesante notar que el problema de evaluación de aproximaciones del **COGP** es también de naturaleza multi-objetivo. Por ello, se recomienda considerar varios indicadores para evaluar los distintos aspectos del desempeño de un algoritmo.

Espaciado ESS(\mathcal{PF}) Propuesto por Schott (1995), el indicador *ESS* (en inglés, "Efficient Set Spacing") mide la desviación de las distancias entre los vecinos más cercanos, en el espacio objetivo. Se considera que mientras menor sea el valor de *ESS*, mejor es la aproximación.

$$ESS = \sqrt{\frac{1}{|\mathcal{PF}|} \sum_{i=1}^{|\mathcal{PF}|} |d_i - \bar{d}|^2} \quad (6)$$

$$d_i = \min_{z^k \in \mathcal{PF}, k \neq i} \|z^i - z^k\| \quad (7)$$

$$\bar{d} = \frac{1}{|\mathcal{PF}|} \sum_{i=1}^{|\mathcal{PF}|} d_i \quad (8)$$

Hipervolumen HV(\mathcal{PF}) Este indicador, propuesto por Zitzler & Thiele (1998) mide el volumen cubierto en el espacio objetivo con respecto a un punto de referencia. Para cada solución $z^i \in \mathcal{PF}$ se construye un hipercubo de volumen v_i . *HV* mide entonces el volumen de la unión de todos los hipercubos: se considera que mientras más grande es el valor del hipervolumen, mejor es la aproximación.

Distancia Máxima DMAX(\mathcal{PF}) Este indicador mide simplemente la distancia máxima entre dos individuos cualesquiera pertenecientes a la aproximación.

$$DMAX(\mathcal{PF}) = \max_{z^k, z^i \in \mathcal{PF}, k \neq i} \|z^i - z^k\| \quad (9)$$

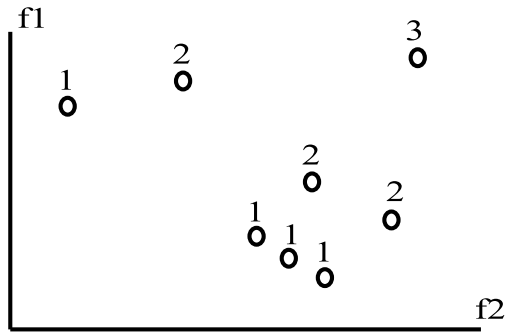


Fig. 1. Ejemplo de cálculo de rango. Algoritmo NSGA-II

3. ALGORITMO NSGA-II

El **NSGA-II** es una versión mejorada del algoritmo **NSGA** original propuesto por Srinivas & Deb (1994). Existen algunas referencias previas relacionadas con la aplicación de este algoritmo para el diseño de estructuras **PID**. Lagunas (2004) utilizó el **NSGA-II** para sintonizar trece controladores robustos de estructura **PID-ISA** de dos grados de libertad. En (Dupuis et al, 2004) se utilizó para diseñar un **PID** para un motor D.C. Los objetivos (*overshoot / settling time / peak time*) se plantearon en función de la diferencia entre la dinámica obtenida durante la simulación y la dinámica deseada. Igualmente, Sánchez et al. (2008) presentaron una comparación entre los resultados obtenidos por **NSGA-II** y **MOGA** con respecto al diseño de dos estrategias de control: **PID** y realimentación de estados.

El algoritmo **NSGA-II** se caracteriza por las siguientes propiedades:

- Utiliza un mecanismo de elitismo.
- Utiliza un mecanismo de preservación de la diversidad de la población, basado en privilegiar a los individuos que habitan en las regiones menos pobladas. Este mecanismo no requiere especificar ningún parámetro para su funcionamiento.
- El mecanismo de ordenamiento es más rápido comparado con otros algoritmos (**MOGA**, por ejemplo).
- Los padres de la nueva generación y los sobrevivientes son seleccionados mediante un torneo binario de tipo $(\mu + \lambda)$.

El mecanismo de selección consiste en asignar un rango a cada individuo de la población. En cada iteración i se determinan los individuos no-dominados y se les asigna el rango i a los mismos. A continuación se eliminan los individuos no-dominados de la población actual y se comienza una nueva iteración (ver figura 1).

Por otro lado, el mecanismo de preservación de diversidad se basa en calcular la suma de las distancias entre individuos adyacentes (*crowding distance*). Entre dos individuos se prefiere al que tenga menor rango. Si los rangos son iguales, entonces se prefiere al que tenga el mayor valor de *crowding distance* (ver figura 2).

4. PROBLEMA DE CONTROL

Los controladores **PID** continúan siendo los más utilizados en la gran mayoría de instalaciones industriales. Su principal ventaja, desde un punto de vista práctico,

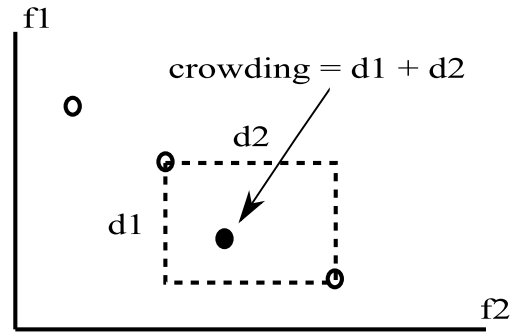


Fig. 2. Ejemplo de cálculo de "crowding distance". Algoritmo NSGA-II

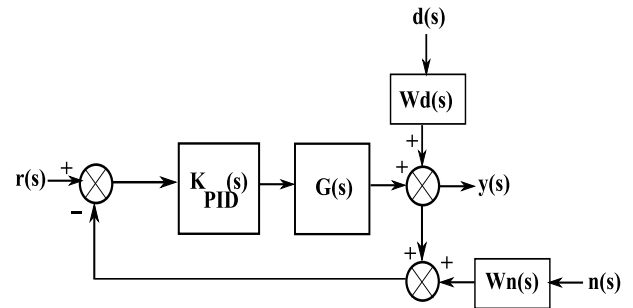


Fig. 3. Lazo de control con controlador PI/PID

consiste en que vienen pre-programados en todos los controladores comerciales (**PLC/DCS**) y son bien conocidos por los técnicos responsables de la instrumentación. De allí la importancia de verificar en primer lugar la posibilidad de resolver un problema de diseño utilizando este tipo de estructuras, antes que ensayar soluciones más complejas.

En este trabajo se considera la estructura definida en (10). Se propone estudiar en particular el lazo de control mostrado en la figura 3. Los bloques $W_n(s)$ y $W_d(s)$ corresponden a funciones arbitrarias de ponderación en frecuencia (filtros).

$$K_{PID}(s) = \frac{K_D s^2 + K_P s + K_I}{s} \quad (10)$$

Las variables que intervienen en el lazo son $r(s)$, $d(s)$ y $n(s)$ como señales de referencia, perturbación y ruido respectivamente. Sin embargo para el problema de regulación se considera $r(s) = 0$. De donde, a lazo cerrado se tiene:

$$y(s) = -\frac{K_{PID}(s)G(s)}{1 + K_{PID}(s)G(s)} W_n(s)n(s) \quad (11)$$

$$+ \frac{1}{1 + K_{PID}(s)G(s)} W_d(s)d(s) \quad (12)$$

$$= -T(s)W_n(s)n(s) + S(s)W_d(s)d(s) \quad (13)$$

con

$$T(s) = \frac{K_{PID}(s)G(s)}{1 + K_{PID}(s)G(s)} \quad (14)$$

$$S(s) = \frac{1}{1 + K_{PID}(s)G(s)} \quad (15)$$

La función de transferencia $S(s)$ se conoce como **función de sensibilidad**, mientras que $T(s)$ se conoce como

función de sensibilidad complementaria. Note que siempre se cumple

$$S(s) + T(s) = 1 \quad (16)$$

A partir de la ecuación (11) se concluye que para compensar los efectos de la perturbación y del ruido, es necesario diseñar el controlador para que las normas de $S(s)$ y $T(s)$ tengan el menor valor posible. Sin embargo, la ecuación (16) implica que no se puede disminuir ambas normas simultáneamente, por cuanto ambos objetivos se encuentran en conflicto. El modelo de planta a lazo abierto viene dado por la siguiente función de transferencia (Hohenbichler & Abel, 2008):

$$G_0(s) = \frac{-0.5s^4 - 7s^3 - 2s + 1}{s^6 + 11s^5 + 46s^4 + 95s^3 + 109s^2 + 74s + 24} \quad (17)$$

y además los filtros W_n y W_d se definen como:

$$W_n(s) = \frac{1}{s+1}, W_d(s) = 1 \quad (18)$$

El problema que se desea resolver en esta sección es el siguiente (note que en este caso la estructura del controlador se encuentra seleccionada de manera *a priori*):

Problema Determine el **COGP** (\mathcal{P}^*) para el problema:

$$\min_{K_{PID} \in \mathcal{K}} f_1(K_{PID}) \quad (19)$$

$$\min_{K_{PID} \in \mathcal{K}} f_2(K_{PID}) \quad (20)$$

con

$$f_1(K_{PID}) = \|W_n(s)S(s)\|_2$$

$$f_2(K_{PID}) = \|W_d(s)T(s)\|_2$$

donde \mathcal{K} es el conjunto de controladores estabilizantes.

5. MÉTODO PROPUESTO

Desde el punto de vista de la ejecución del algoritmo genético, para el inicio del proceso de búsqueda se debe establecer el espacio de variación de los parámetros K_P, K_D, K_I . En el caso general puede ser muy difícil generar y preservar individuos estables y/o cercanos al frente de Pareto. Recientemente, Hohenbichler & Abel (2008) presentaron el toolbox **PIDROBUST** el cual permite calcular los límites de la región de controladores estabilizantes en el espacio (K_P, K_D, K_I) .

La idea básica para este cálculo se basa en el teorema presentado por Ho et al. (1998) en el cual se demuestra que las regiones estabilizantes en el plano (K_D, K_I) para un K_P determinado, están formadas por polígonos convexos cuyos límites pueden ser determinados. Hohenbichler & Abel (2008) resuelven entonces el problema del cálculo del intervalo para K_P de manera que exista al menos un plano (K_D, K_I) estabilizante, **en el caso de sistemas lineales escalares que no posean ceros en el eje imaginario.**

En el caso del modelo de planta planteado en la sección anterior, el toolbox **PIDROBUST** permite determinar los límites del espacio factible, tal como se muestra en la figura 4.

Durante la primera etapa del método de diseño se utiliza este toolbox para encontrar los límites de la región de

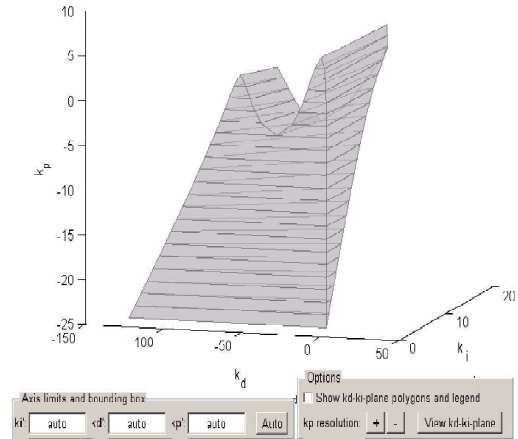


Fig. 4. Espacio de controladores estabilizantes para la planta G_0 definida en la sección 4.

controladores estabilizantes. En la figura 5 se muestra un ejemplo de los polígonos que conforman el espacio de controladores estabilizantes para un valor de $K_P = 1$.

Durante la segunda etapa se utiliza el **NSGA-II** para resolver el problema de diseño propiamente dicho, tomando en cuenta los límites obtenidos durante la primera etapa (ver figura 6). Los operadores de mutación y cruce fueron adaptados para que los individuos generados se mantengan al interior del espacio factible.

Para tener una idea de la eficiencia del método obtenido, se intentó resolver el mismo problema utilizando un enfoque tradicional que consiste en generar las soluciones iniciales de manera aleatoria y uniforme en el espacio

$$K_P, K_D, K_I \in [-200, 200] \quad (21)$$

y utilizando una función de penalidad de la forma:

$$p_i(x) = \begin{cases} f_i(x) & \text{si } \max \{ \text{eig}(S(x)) \} < -\delta \\ \eta * \max \{ |\text{eig}(S(x))| \} + M, & \text{en caso contrario} \end{cases} \quad (22)$$

con $\delta = 10^{-4}, \eta = 10^3, M = 10^3$. Note que la selección de los límites favorece al método tradicional puesto que encierra de manera ajustada la región de controladores estabilizantes, aún cuando esta información no se dispone *a priori*.

6. RESULTADOS

El método propuesto se puso en práctica codificando los algoritmos en **MATLAB**[®], versión 7.0. Para el **NSGA-II** se utilizó la configuración mostrada en la tabla 1. En la figura 7 se muestra un ejemplo del tipo de aproximaciones que es posible obtener. El eje de las abcisas corresponde a los valores de $f_1 = \|W_n(s)S(s)\|_2$. El eje de las ordenadas corresponde a $f_2 = \|W_d(s)T(s)\|_2$. En la tabla 2 se muestran los resultados obtenidos para los tres indicadores, en promedio y desviación estandar sobre 30 ejecuciones. Para el indicador de hipervolumen (*HV*) el punto de referencia considerado fue $(1, 1)$.

En cuanto a los resultados obtenidos mediante el método tradicional, luego de 30 ejecuciones, cada una de 50 generaciones, no se logró conseguir ninguna solución factible.

