Daily Generation Management at Electricité de France: From Planning Towards Real Time

Arnaud Renaud

Abstract—This paper is intended to show the actual interest the application of new optimization methods may have in order to improve the operation of power systems.

With this aim, the work carried out over the last few years to optimize day-to-day operation of electricity generation is presented.

First, we offer a critical analysis of the modeling and methods implemented today for daily operation.

Second, we present the methodological breakthroughs which have been achieved in the last two years.

Finally, we show how these breakthroughs will be used:

- in a new tool for optimizing generation schedules—the APOGEE program,
- in on-line optimization functions for control centers.

I. INTRODUCTION

THE Electricité de France generation mix comprises some 60 nuclear power plants, about a hundred conventional thermal power plants and more than fifteen valleys (hundreds of hydro plants). Over and above its sheer size, the structure of the mix has two specific characteristics.

First, generation costs differ greatly from one facility to the next. The nuclear fuel is much cheaper than fossil fuels: the choice of one type of power plant rather than another is far from innocent since major economic implications are involved.

Besides, in a given year, over 90% of french electricity is generated by the nuclear (75%) and hydraulic power plants, facilities which "memorize" their operation. This is particularly obvious in the case of hydraulic power facilities, since operation in this case concerns interconnected reservoirs of energy (i.e., water). Given the high dispersion in the ratio between the reservoir content and power output capacity, the stock constraints appear at every time scale.

The core of the nuclear power plants is refueled once every year (or every other year), and the refueling operation along with the various maintenance operations it implies requires the plant to be shut down for several weeks. The stock of fuel must therefore be managed on

an annual basis. At the same time, the number of changes in power output is limited on a daily basis, particularly in order to limit the production of radioactive waste. As a result, there are also "controls supplies" which have to be managed on this time scale.

To these two specific characteristics is to be added the more conventional factor (since it has to be dealt with by all companies generating electricity), linked to start-up operation. The start-up costs for thermal power plants and the nonconvexity of their operation domain (when the plant is on-line, the minimum level of power output cannot be zero) are both factors which have to be taken into account.

Consequently, it is necessary to take into account the economic side of the problem (the different generation costs) in addition to the spatial (the number of power plants to manage) and temporal (major dynamic constraints) complexity. These simple remarks suffice to prove that real-time control of the system of electricity generation can only be performed under acceptable conditions if the corresponding decisions are fully prepared in advance.

Today this preparation is broken down into two phases.

Mid-Term Operation: Investigates a time scale ranging from a few months to several years. On the basis of historical data concerning weather temperatures and consumption forecast estimates, strategic decisions pertaining to power output operation, such as choosing nuclear power plant shutdown dates, have been developed.

With regard to the operation of water reservoirs and nuclear fuels, the solution which consists in establishing a policy, thus creating constraints for shorter term optimization functions, (quantity hierarchization) has been rejected, in preference for a form of "price hierarchization."

This choice is essentially the consequence of the high variability of the water inflows and of the demand. The mid-term operation indicators are only revised once or twice a month. Therefore, drawing up a policy of water discharge and of nuclear generation would lead to a nonoptimal or even unfeasible solution.

Mid-term operation determines Bellman values. These enable estimates of the values for each of the major energy reserves (large hydraulic water reservoirs and the cores of nuclear plants) according to their current levels (see Fig. 1).
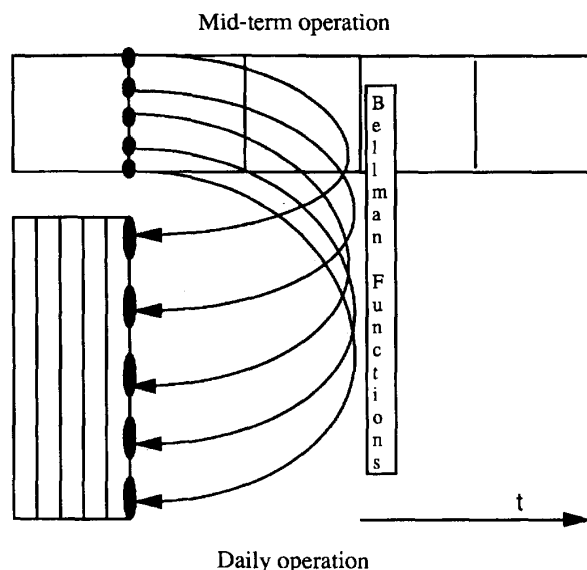
Mid-term operation

Daily operation

Fig. 1. Mid-term operation generates Bellman functions.

Daily Operation: Determines on the basis of indications developed at mid-term operation level, the day to day instructions to be given enabling real-time operation with security and economy.

Considering the complexity and the size of the problems which are related to the mid-term and daily operation at Electricité de France, the integration of all the dimensions of these problems appears to be particularly difficult. In fact, the optimization softwares which are in use today do not cope with an exhaustive modeling of the generation-transmission system. These softwares are therefore used in a study mode: their results are not directly implemented but first of all examined by one or several experts and they are often run several times before the corresponding decisions are made.

Nevertheless, the advantage of having tools resembling automata appears greater, the closer one gets to considering real-time problems. Furthermore, a real optimization of the generation schedules during the day will soon be possible owing to the quick improvement of processing and measuring capacity.

It is for this reason that we have chosen in this paper to focus more specifically on this problem. The gradual advance towards real time offers the opportunity of querying the type of modeling used up to the present date on the problem. Moreover, on-line use of a model requires accuracy and efficiency. Such a perspective thus provides a powerful incentive to search for new numerical methods.

The problem of day-to-day operation is considered at the present time as a question of planning. Each day, generation schedules are drawn up for the next.

The problem which then has to be solved thus becomes one of large-scale optimization of a composite nature (cost and constraint functions). At the EDF, this problem is solved by means of the OCTAVE software system which has been in use for the last seven years.

The choices pertaining to the modeling and resolution methods which were made at the design stage of the software are presented below.

Second, we shall be considering within this context the methodological improvements which have been achieved over the last two years.

Finally, we present an overview of the new perspectives offered by the upgrading of the hardware system which has been implemented at the national control center level. The new system enables access to and display of the requisite data for an on-line generation scheduling optimization. By looking at the difficulties created by the upgrading mentioned above, we then show that the recent methodological developments offer major advantages.

II. TODAY'S DAILY OPERATION PROCEDURES

A. Modeling

A.1) Drawing up Generation Schedule the Previous Day

From one day to the next, optimizing the generation of electricity appears to be a stochastic problem. Two main sources of random factors can be distinguished.

First, forecasting consumers demand cannot be exact. Nevertheless, the approximation related to the modeling of the consumer behavior is not the main cause of uncertainty: the uncertainty linked to weather temperature forecasts has, in winter time, a major effect. Indeed, the use of electric heating leads to a variation of more than 1000 MW in demand, for a variation of one Celsius degree in the outside temperature (in comparison with peak power output levels of roughly 60000 MW). As a result, since the temperature cannot be forecast on a daily time scale to more than within a few degrees, uncertainty in terms of demand is a factor which cannot be neglected.

Parallel to this, the outage rate for a given thermal power plant is of the order of 10^{-3} per hour. For this reason, the loss of one plant, and possibly two, has to be taken into account on the daily time scale.

The problem is written, however, within a deterministic framework, and to do this, a notion of reserve is introduced. This power margin is defined as being the power which can be produced in less than twenty minutes over the whole of the generation mix. Provided that an "adequate" level of reserve is available, the operator can meet most of the random occurrences on a daily basis. Therefore the day to day operation of generating electricity is "reduced" to the problem of minimizing generation costs matching forecast demand and providing a sufficient power margin.

Formally, the corresponding optimization problem can be expressed as follows:

$$\min_{p_i, \mathscr{T} \in \mathscr{P}_i} \sum_{i \in \mathscr{I}} c_i(p_i, \mathscr{T}) \qquad (1)$$

subject to:

$$\sum_{i \in \mathscr{I}} p_{i,t} = D_t \qquad \forall t \in \mathscr{T} \qquad (2)$$

$$\sum_{i \in \mathscr{I}} r_{i,t} \geq R_t \qquad \forall t \in \mathscr{T}, \qquad (3)$$

where

- if $e \in \mathbb{R}$ and $\mathscr{F}$ is a set, $e_{\mathscr{F}}$ denotes the vector $(e_f)_{f \in \mathscr{F}}$,
- $\mathscr{I}$ is the set of the generation units (thermal plants and valleys),
- $\mathscr{T}$ is the period under study (the following day),
- $p_{i,t}$ is the generation of the plant i at time t and $p_{i,\mathscr{T}} = (p_{i,t})_{t \in \mathscr{T}}$ is the vector representing the output of power plant i during the day,
- c_i is the operating cost function,
- D_t is the demand and R_t the reserve requirement at time t,
- $\mathscr{P}_i$ is the operating domain of the unit i (if $p_{i,\mathscr{T}} \in \mathscr{P}_i$ then the generation schedule $p_{i,\mathscr{T}}$ respects the operating dynamic constraints of unit i),
- $r_{i,t}$ is the reserve that can be supplied by unit i at time t, given that its generation schedule is $p_{i,\mathscr{T}}$.

Most electric utilities optimize thermal power plant start-ups on an open-loop basis ("Unit Commitment") according to a similar principle. The dispatcher is thus capable of meeting daily random occurrences thanks to the reserve provided by power plants which are up but not at their maximum power output. Therefore, this formalization leads to a hierarchization which supposes that random factors are of a sufficiently moderate amplitude for start-up decisions not to be updated, other than in exceptional circumstances, during the day.

The hypothesis underlying this is quite similar. What we are looking for, however, are generation schedules and not just start-up controls. This point is justified by the fact that the operation of the Electricité de France generation mix, as we have already emphasized, involves the handling of dynamic constraints (stock constraints). As a result, even if power plant generation schedules have to be partially updated during the day, these constraints cannot be neglected when drawing up the schedules: the hypothesis which consists in relaxing the dynamic constraints, otherwise those linked to the start-up would lead to too great an approximation. For a valley, for example, the reservoir content bounds impose the tightest constraints, not the start-ups.

In this way the detailed representation of operating constraints for power plants can, at least initially, be understood as being similar to the conventional lengthening of the periods under study in order to avoid boundary effects: the solution of the problem (1) will also be partially updated but the modeling for power plants is sufficiently fine to avoid any discontinuity between the optimization and the "real" operating constraints.

There is another cause, however, for this detailed representation. While there exists some uncertainty about consumer demand levels and the availability of the generation units, it only concerns a few percent of demand (a few thousands of MW in relation to some 60,000). Consequently, the overriding majority of power output schedules obtained the previous day by solving (1) will in fact

be implemented, as well as being fully justified both technically and economically.

As a result, generation plans can be sent the previous day to the power plants. It enables their feasibility to be verified at local level.

Furthermore, the problem then faced by the dispatcher, in real time, will be relatively small. Basically he will only have to consider the power plants contributing to the reserve power. Thus he will only have to manage the discrepancies between forecast and real demand, a problem which is incomparably simpler than that posed by the global optimization of daily generation schedules and which can be effectively controlled even in real time conditions.

Therefore, this formalization yields to an aggregated real time problem. This implicit aggregation enables the flow of information to be exchanged during the day to be limited. The operator only has to send a few instructions to a reduced number of power plants. Since order transmission procedures are at the present date not very computerized, the advantage of such a procedure is quite obvious.

The arguments we have put forward so far show the interest in having a quality generation policy available the previous day. To my mind, this interest explains to a large extent the formalization (1) which, for more than five years, has been the basis for daily generation planning at EDF. The arguments, however, in no way justify the deterministic modeling which has been retained here. Using stochastic optimization could well supply a generation policy and even feedbacks which could be used as indicators as part of the daily operation process.

Here the obstacle becomes one of methodology. Stochastic control theory fitted to large-scale systems, for the most part, remains to be developed. And in addition to the difficulty of the principle itself, there must also be taken into consideration the fact that the problem posed is of a particularly large scale. Therefore, taking into account "real" problem's stochastic dimensions together with a detailed representation of the operating constraints must be considered as a long term perspectives.

It should be emphasized that even when we fall back to the deterministic framework, solving the problem poses highly complex technical difficulties. Even today, it is not yet possible to handle all the constraints in the global optimization. We shall see, for example, that the operation of valleys has to be performed in two stages.

Similarly, one might be surprised by the fact that the constraints related to the transmission system do not appear in (1). Compared to the construction costs of the transmission lines the generation investments are very heavy. This fact leads to the construction of a large-scale grid system in relation to output capacities: power plants thus have only limited constraints on the daily time-scale.

These constraints do exist however, and in practical terms often result in limiting the operation of certain plants.

It therefore appears that the formalization (1) is the

result of a necessary compromise taking into account the data processing capacities and the performance of numerical methods available to us. Since improvements can be expected in both fields, we shall see how we should be able to enrich the modeling proposed herein in the mid and long terms.

A.2) Optimizing with Half-Hour Steps

The period under study is discretized into 48 steps of half an hour each. In practice, discrete steps of this size seem to be sufficiently fine given the objectives of the optimization which is performed: various means of adjustment (secondary frequency control, primary control), the details of which we do not go into here, enable the adjustment of output according to very short-term consumption demand (over a few minutes).

The operating constraints which are further detailed have no meaning other than in this context.

A.3) Operating Constraints and Costs

Nuclear and Thermal Power Plants Operating Constraints:

In order to be considered feasible, the daily generation schedule of a nuclear power plant must observe the following constraints:

C1) the minimum interval between changes in generation must be at least two hours,

C2) certain ranges of output are prohibited,

C3) the number of changes in output instructions is limited to four per day,

C4) only one major drop in output is authorized per day (by major drop is meant the switch to a setpoint generation near that of the minimum generation).

While all the constraints have to be taken into account they do not all have the same meaning. Constraint C1) and C2) aim at guaranteeing operational safety. On the other hand, while major drops and the number of changes in instructions are limited over life time of the nuclear plants, this is not the case for the daily time scale. These constraints therefore correspond to a form of "quantity hierarchization" which is justified by the fact that the constraints introduced only imply low cost.

On the other hand, the generation schedules of thermal power plants have to observe the following constraints:

C1)' if the plant is on-line, $p_{i,t} \in [\underline{p_i}, \overline{p_i}]$ with $\underline{p_i} > 0$,

C2)' a plant must not have more than one start-up per day,

C3)' during shutdown or start-up phases, typical output curves must be followed,

C4)' minimum down time.

In both cases, operating ranges are nonconvex. Furthermore, the discontinuity has too great an amplitude to not be modelized. The operation of start-ups for thermal power plants is in fact one of the prime objectives of the procedures described in the present document.

In consequence it is vital that the impossibility of generating below a certain minimum power output $(\underline{p_i})$ be taken into account.

Nonconvexities are therefore not eliminated. They are even exaggerated by only authorizing a finite number of operating points (when only power output ranges are prohibited in steady state). This attitude which *a priori* may seem paradoxical is justified by the fact that, as we shall see later, these constraints have given rise to resolutions by dynamic programming, a method whose implementation is rendered easier by this modeling.

Generation Costs: The operating costs for a thermal power plant (conventional or nuclear) when started-up is modelled as an affine function of power output level. During shutdown the cost is presumed to be zero. To the start-up controls of conventional power plants (the reader is reminded that nuclear power plant "start-ups" are not managed at a daily level) is linked a cost.

The Hydro Systems:

Constraints: A valley consists in a series of interconnected reservoirs. After a fixed period for its course, the discharged water reaches the downstream reservoir.

Moreover, some hydro plants can pump (consumption of electricity) water from a reserve to the upstream reserve during periods of low demand, water which is discharged (generation) when generation costs become higher again.

Formally, operating constraints to be taken into account are expressed as follows:

$$\forall h \in \mathscr{H}_v, \qquad \forall t \in \mathscr{T}:$$
$$V_{h,t+1} - V_{h,t} = a_{h,t} - T_{h,t} + Q_{h,t} - D_{h,t}$$
$$+ \sum_{g \in \Gamma(h \leftarrow)} \left(T_{g,t_{g \to h}} - Q_{g,t_{g \to h}} + D_{g,t_{g \to h}} \right)$$
$$V_{h,t} \in \left[\underline{V_h}, \overline{V_h} \right], \qquad D_{h,t} \in \left[0, \overline{D_h} \right],$$
$$T_{h,t} \in \left[\underline{T_{h,t}}, \overline{T_{h,t}} \right], \qquad Q_{h,t} \in \left[\underline{Q_{h,t}}, \overline{Q_{h,t}} \right],$$
$$(4)$$

where

- $\mathscr{H}_v$ is the set of water reservoirs for the valley under consideration—each reservoir is related to a plant having the same index,
- $\Gamma(h \leftarrow)$ is the set of plants upstream of reserve h,
- $V_{h,t}$ is the content and $a_{h,t}$ the inflow of the reservoir h at time t,
- $T_{h,t}$ is the discharge, $D_{h,t}$ the spillage and $Q_{h,t}$ the quantity of water pumped by plant h and time t,
- $t_{g \to h} \stackrel{\text{def}}{=} t - d(g \to h)$ with $d(g \to h)$ denoting the delay for discharge of plant g to reach reserve h,

This modeling neglects two types of constraints:

Spillage Constraints: Generally, spillage is not possible if the reservoir is not full. Therefore constraints of the type $(\overline{V_h} - V_{h,t}) D_{h,t} = 0$ should be handled.

Environmental Constraints: The discharge variations must be smooth in order to ensure the river flow to be steady.

Not being able to handle these constraints in the global optimization program which is sketched hereafter, a heuristic is used to find feasible solutions. Following the course of the river, the generation schedule of each plant

is established using dynamic programming. Taking into account the discharge and the spillage of upstream reservoirs, this local optimization processes the feasible planning which is the closest from the solution of the global optimization.

In practice, this heuristic yields satisfactory results: indeed spillage leading to a reduction in stock without generation are in fact very rare in schedules established the previous day for the next.

Reservoir Operation: As already emphasized, the largest water stocks are quantified by mid-term generation operation programs. And the Bellman functions obtained are related only to the level of each stock.

The cost function taken into account in daily operation is the derivative of these Bellman values around the current level. This approximation appears to be satisfactory considering the small level variation of these reservoirs during a day.

Smaller water reserves are not explicitly taken into account by mid-term stock operation. These reservoir contents must always be higher than a reference level at the end of the day.

Generation: Generation is considered to be a concave function of discharge.

In fact the efficiency/discharge relation is also parameterized by the plant head (the difference between the forebay and tailwater elevations). Indeed, the higher the level of water, the higher the output efficiency.

In France, however, daily generation operation ignores this factor. During a day, the volume variation of the largest stocks (whose water is valorized by mid-term operation) is not significant. Furthermore, the operation mode for smaller reserves results in the water level being maintained near a reference value. In this way, on a given day, the stake corresponding to the operation of the plant head on the daily time-scale is not very important.

Furthermore, the generation is considered to be a piecewise-linear function of the discharge in order to enable the use of highly efficient linear programming methods.

B. Resolution Method

B.1) Price Decomposition

The method of resolution retained is based on a price decomposition [11]. The coupling constraints (2) and (3) are dualized and the "prices" updated by a subgradient algorithm.

The method is used today by a majority of electric utilities to process this type of problem ("unit commitment").

First of all, the choice of this method is justified by the scale of the problem to be solved. It should be remembered that in France generation schedules for more than one hundred thermal power plants and more than fifteen valleys have to be optimized over a twenty-four-hour period discretized into half hour steps. Given the dynamic constraints which have to be taken into account, decomposition has an obvious advantage. Its benefit

is even greater if we take into consideration the high level of heterogeneity of the generation facilities. The decomposition strategy—and this is one of its major advantages—actually enables the method of resolution to be tailored to the type of subproblem being processed. In this way, the hydraulic subproblems can be solved by linear programming, while the thermal subproblems are processed by dynamic programming.

Given the decomposition principle, we may then investigate the choice of a decomposition mode. In the problem of generation operation we are dealing with, cost function (start-up costs) has poor mathematical properties (the derivatives are not Lipschitz). In this context, price decomposition seems particularly well suited. Quantity or prediction decomposition methods require more restrictive assumptions to be verified (in particular the existence and Lipschitz property of the derivative of the cost function).

The convergence of price decomposition methods does not require such assumptions [7]. In addition, even if the domain is not convex (which as we shall see is the case here), the maximization of the dual function yields an approximation of the optimal solution.

Parallel to this methodological justification we might also underline the interest for the engineer of such a method. The decomposition principle and the local processing of operating constraints which results from it, enables these constraints to be upgraded without requiring to update the whole of the optimization program: only local modifications to the software have to be performed. In this way, methods based on decomposition principle seem much more open-ended, and therefore much easier to use in practice than the conventional large-scale problem methods (linear programming for example).

B.2) Coordination: The Uzawa Algorithm

Having selected a price decomposition principle, we use the Uzawa algorithm [1] which may be considered as a gradient (or subgradient) algorithm to maximize the dual function. Because the cost function is separable, this algorithm allows decomposition and has the following structure (iteration $k + 1$):

$$\forall i \in \mathscr{I}: \quad \begin{array}{c} \min_{p_{i,\mathscr{T}} \in \mathscr{P}_i} c(p_{i,\mathscr{T}}) - \sum_{t \in \mathscr{T}} \lambda_t^k p_{i,t} - \sum_{t \in \mathscr{T}} \mu_t^t r_{i,t} \\ \Downarrow \\ p_{i,\mathscr{T}}^{k+1}, r_{i,\mathscr{T}}^{k+1}; \end{array}$$

$$\forall t \in \mathscr{T}: \quad \begin{array}{c} \lambda_t^{k+1} = \lambda_t^k + \epsilon_p^k \left(D_t - \sum_{i \in \mathscr{I}} p_{i,t}^{k+1} \right), \\ \mu_t^{k+1} = \max \left\{ \mu_t^k + \epsilon_r^k \left(R_t - \sum_{i \in \mathscr{I}} r_{i,t}^{k+1} \right), 0 \right\}; \end{array} \quad (5)$$

where $\forall k \in \mathbb{N}$, $\epsilon_p^k > 0$ and $\epsilon_r^k > 0$.

The original problem is therefore replaced by a sequence of subproblems related to each of the generating units: handling its own operating ranges, each of these units minimizes, at each iteration, a balance taking into

account its production cost and the "revenue" from its share in generation and reserve.

Since the cost function is not strictly convex, the dual function is not necessarily differentiable. We therefore have to use a subgradient algorithm.

Convergence to a maximum of the dual function is ensured if the sequences $(\epsilon_p^k)_{k \in \mathbb{N}}$ and $(\epsilon_r^k)_{k \in \mathbb{N}}$ are two sequences of type σ.—The reader is reminded that a sequence $(\epsilon^k)_{k \in \mathbb{N}}$ is a sequence of type σ if $\sum_{k=0}^{k=+\infty} \epsilon^k = +\infty$ and if $\sum_{k=0}^{k=+\infty} (\epsilon^k)^2 < +\infty$—.

If, under these assumptions, convergence to a maximum of the dual function has been proven, it is not the case for the primal. We have already emphasized that the operating range for thermal power plants defined a nonconvex set. In this context, since the existence of a saddle point is more than uncertain, it seems to be quite difficult to ensure the convergence of the Uzawa algorithm towards such a point.

We shall return in more details to the problems of convergence posed by this algorithm in paragraph A.1).

Otherwise, decomposition does not ensure the CPU time to be short enough. The generation mix includes more than one hundred plants and the subgradient algorithm used requires a hundred iterations. As a result, there are more than ten thousand local problems which have to be processed. Considering the relative complexity of the operating constraints which have to be taken into account, it becomes quite clear that well-suited, efficient methods have to be sought to solve the local problems.

The paragraph which follows outlines the methods used to solve the local hydraulic problem.

To avoid useless repetition the examination of the difficulties posed by the resolution of the thermal subproblems and the solutions we have been able to bring to them, will be outlined in the paragraph on the new software system APOGEE.

B.3) The Hydro Subproblems: An Efficient Implementation of the Simplex Method

The greatest problems of efficiency are met in the processing of this problem. The hydro-valleys can include more than ten interconnected reservoirs. In this case, 480 constraints (4) have to be handled (48 time steps and ten water reservoirs).

Consequently, the "local" hydraulic problems have a quite an appreciable size. Their method of resolution must therefore take full advantage of the specific structure of the problem.

Within this structure, one point in particular is worth noting: the constraint matrix, related to the constraints (4) and which we will denote by A, is the incidence matrix of a graph G from which a node has been removed (see Fig. 2).

Solving the hydrosubproblems is therefore to seek a flow of minimum cost. In this framework it is possible to make highly efficient use of the simplex algorithm [10].

Let I_k be the basis of the k iteration and A^{I_k} the basis matrix and assume that A has $m - 1$ rows (m is there-
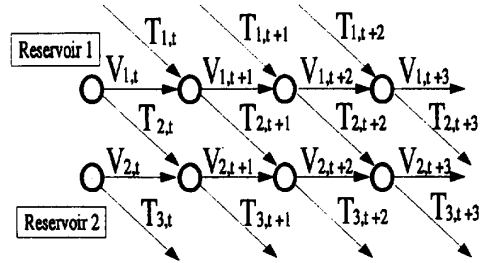

Fig. 2. The constraint matrix has a flow structure.

fore the number of nodes of the graph G). The set of arcs related to the columns of the basis matrix is a tree of the graph G (see Fig. 3).

Solving this problem by the simplex algorithm, the main difficulty consists in the resolution of systems $\lambda A^{I_k} = c$ (computation of the relative cost coefficients) and $A^{I_k}x = b$ (computation of the displacement direction of the basic variables). Nevertheless, solving these problems is obvious (and does not need more than $O(m)$ operations) providing that are known a "numbering" and an "orientation" functions (respectively, σ_k: $\{1,\cdots, m - 1\} \to \{1,\cdots, n\}$ (n denoting the number of arc of G) and ϵ_k: $\{1,\cdots, m - 1\} \to \{-1, 1\}$) which have the following properties:

• removing the arc $\sigma_k(i)$ ($i \in \{1,\cdots, m - 1\}$), the basis tree yields two trees. The arc numbers of one of those trees are all lower than i,
• the arc i such that $\sigma_k(m) = i$ is adjacent to the artificial node,
• if $\epsilon_k(i) = 1$ the arc i is, in the tree, oriented towards the artificial node.

Assume such functions σ_k and ϵ_k to be known at the kth iteration, we will see how to update these functions. First of all, it may be noticed that adding an entering arc (an entering column) to the basis tree yields a cycle $\mathscr{C}_k$. And an arc of this cycle has to be removed to generate a new basis tree. It may be proven that the cycle $\mathscr{C}_k$ is the set of arcs related to the nonzero components of the solution of the system $A^{I_k}x = A^e$ where A^e is the entering column.

Having solved this system and having found the leaving and the entering arcs, σ_k and ϵ_k are updated as follows:

• the numbering and the orientation of the arcs whose numbers are greater than n_s (n_s denoting the number of the arc leaving the basis) do not change,
• from 1 to $n_s - 1$ the arcs which do not belong to $\mathscr{C}_k$ are renumbered in an increasing order; their orientation is not modified ($\epsilon_{k+1}(i) = \epsilon_k(i)$),
• from the node of the arc s which is the furthest from the artificial node to the arc e, the arcs are numbered (still in an increasing order); the orientation of these arcs is changed ($\epsilon_{k+1}(i) = -\epsilon_k(i)$),
• the arc e is numbered; assuming $e = (x_1, x_2)$, x_1 and x_2 being two nodes of the graph G, if x_1 is adjacent to the chain joining s to e then $\epsilon_{k+1}(e) = -1$ and if x_2 is adjacent to this chain $\epsilon_{k+1}(e) = 1$,
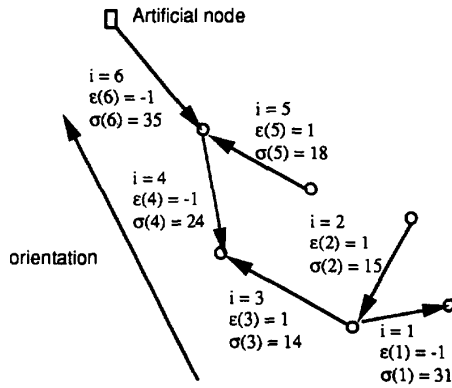
Fig. 3. The basis is a tree.

Fig. 4. The dual functions ϕ and ϕ_c.

• the remaining arcs of $\mathscr{C}_k$ are then renumbered (following the order defined by σ_k), the signs of these arcs are not modified.

Parallel to this, the linear program used for this resolution makes use of a technique known as "multiple pricing."

For a given iteration, the set of variables which can enter the basis is memorized. The iterations of the simplex algorithm which follow will only accept candidates belonging to this set until there are none left. At this point the whole set of candidates is reinitialized. This method, which can be seen as an aggregation method within the variable space, enables savings of between 30–50% in computing time.

Globally, the time required for a resolution by this process is 50 to 200 times shorter than that which would result from the use of the standard simplex algorithm. On an IBM 3090 one resolution of the fifteen subproblems related to the national valleys will only take about twenty seconds.

In addition, at each iteration, only the dual variables associated with the coupling constraints change. As a result, the solution obtained at the previous iteration is always a feasible basic solution. What is more, the dual variables evolution being rather slow during the iterations, the initialization will be very accurate.

In this way, the optimal basis related to each valley is stored at each iteration and used to initialize the following iteration.
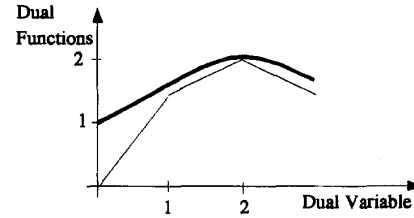
This process, in practice, turns out to be remarkably efficient. A global computing time of one minute is enough on an IBM 3090 (or approximately 5 minutes on a SUN 4 2/75 workstation), for one hundred iterations. As a result, one hundred iterations require no more than three times the computing time for the first one.

III. CURRENT DEVELOPMENTS AND PERSPECTIVES

A. Methodological Breakthroughs

A.1) The Limits of Conventional Duality

We have already underlined the fact that the modeling of thermal power plants defines nonconvex operating domains. As a result, the optimization problem we have to solve does not admit any saddle-point and there can therefore be no hope of exact convergence: the coupling constraints on demand and on the power reserve cannot be met exactly.

However, difficulties resulting from the weakly convex nature of the cost function exist in addition to those linked to the nonconvexity of the problem.

As a result, even in the convex case (i.e., a convex cost function and domain), the differentiability of the dual function cannot be ensured. To reach this conclusion there is no need to consider complex problems.

The following problem may be related to the optimization of the generation of a two plants mix (see Fig. 4):

$$\min_{p_1, p_2} p_1 + 2p_2 \qquad (6)$$

subject to

$$p_1 \in [0,1], \qquad p_2 \in [0,1],$$

$$p_1 + p_2 = 1.5. \qquad (7)$$

Dealing with the constraint (7) by price decomposition leads to the definition of the Lagrangian $\mathscr{L}$: $(p_1, p_2, \lambda) \mapsto 1.p_1 + 2.p_2 + \lambda(1.5 - p_1 - p_2)$.

The corresponding dual function ϕ is defined by the equality:

$$\phi(\lambda) = \min_{p_1 \in [0,1]\, p_2 \in [0,1]} p_1 + 2.p_2 + \lambda(1.5 - p_1 - p_2) \qquad (8)$$

The dual function being nondifferentiable the use of a subgradient algorithm is required. Consequently, the convergence is relatively slow. This, however, is not the main difficulty.

Let us suppose that we implement the Uzawa algorithm to maximize ϕ. After a certain number of iterations we would find ourselves near the solution of the dual problem $\lambda = 2$. But what about the primal?

For an iteration k, if λ_k is greater than 2, (8) will yield ($p_1 = 1$, $p_2 = 1$). And if the value is less than 2, we shall find ($p_1 = 1$, $p_2 = 0$). Thus, the optimal solution ($p_1 = 1$, $p_2 = (1/2)$) cannot be obtained and the coupling constraint (7) will not be met.

This difficulty can be attributed to the fact that any element of $\{1\} \times [0,1]$ is a solution of the minimization problem which defines the dual function (8). And if the

set $\{1\} \times [0,1] \times \{2\}$ contains the saddle point $(1,(1/2),2)$, knowing this does not entail it can be determined.

From a more formal point of view, then, the Lagrangian is not "stable" [7]. In the convex case, the stability of the Lagrangian can only be proven, within the framework of the theory of "classic" duality, if the function to be minimized is strictly convex. Broadly speaking, the operating costs of each of the power plants are affine functions of outputs. In this way, the stability of the Lagrangian related to our problems of daily generation scheduling (even convex ones) cannot be ensured.

The theory proposes a solution to these difficulties: the Augmented Lagrangian.

A.2) The Augmented Lagrangian

Consider the following problem:

$$\min_{u \in U^{\text{ad}}} J(u) \qquad (9)$$

subject to

$$\Theta(u) = 0$$

where

- U^{ad} is a convex closed subset of an Hilbert space $\mathcal{U}$,
- $J: \mathcal{U} \to \mathbb{R}$ is a convex function and $\Theta: \mathcal{U} \to \mathcal{C}$ is a continuous affine mapping.

The Augmented Lagrangian mapping [12], corresponding to this problem ($\mathcal{L}_c = \mathcal{U} \times \mathcal{C}^* \to \mathbb{R}$ [$\mathcal{C}^*$ being the dual of $\mathcal{C}$]), is defined as follows:

$$\forall \lambda \in \mathcal{C}^*, \qquad \forall u \in \mathcal{U}:$$

$$\mathcal{L}_c(u, \lambda) = J(u) + \langle \lambda, \Theta(u) \rangle + \frac{c}{2} \|\Theta(u)\|^2, \quad (10)$$

where $\langle \cdot, \cdot \rangle$ is the scalar product.

This Lagrangian, like the ordinary Lagrangian, is a convex function in u and concave in λ. Moreover, in the convex case, the saddle-points of these Lagrangians are the same.

This result can be easily understood $\mathcal{L}_c$ being the sum of an ordinary Lagrangian and a penalty term. The Augmented Lagrangian thus appears to result from the combination of price decomposition and penalty techniques.

This interpretation, however, is perhaps not the most useful: this technique may be seen as resulting from an operation of regularization. Indeed, it can be shown that the dual function ϕ_c related to $\mathcal{L}_c$ is such that: $\forall \lambda \in \mathcal{C}^*$: $\phi_c(\lambda) = \sup_{\mu \in \mathcal{C}} (\phi(\mu) - (1/2c)\|\lambda - \mu\|^2)$. Thus, the dual function ϕ_c is the Moreau transform of ϕ. Now this transformation, when applied to a concave function, has two remarkable properties:

- the maximization of the regularized function is equivalent to the maximization of the original one,
- the regularized function is differentiable with Lipschitz derivatives.

If we return to example (6), Fig. 4 enables us to illus-

trate the effect this regularization operation has on the dual function.

Moreover, having found a "right" price, a dual variable maximizing the regularized dual function, the primal minimization of the Augmented Lagrangian yields (in the convex case) a solution of the original problem.

Therefore, this technique enables all the difficulties above-mentioned to be overcome—except those related to the nonconvexity—. It does, however, have the major inconvenience of introducing nonseparable terms. Thus the Usawa algorithm, for example, will not enable the decomposition which justifies the choice of dual methods.

The Auxiliary Problem Principle copes with this new difficulty.

A.3) Decomposition and Augmented Lagrangian

Amongst the best-known algorithms to seek saddle-points of a Lagrangian stand those of Uzawa and Arrow–Hurwicz [1].

The Uzawa algorithm, as we have already seen, is a gradient (or subgradient) algorithm applied to the maximization of the dual function. In practice it leads to alternating an update, by a gradient step, of dual variables and an explicit resolution of the minimization problem of the Lagrangian.

The Arrow–Hurwicz algorithm avoids the minimization phase of the Lagrangian by alternating a gradient step on the primal and dual spaces.

Neither of these algorithms can be used to implement an Augmented Lagrangian technique to optimize daily generation schedules. The Uzawa algorithm, as we have already emphasized, does not enable decomposition. That of Arrow–Hurwicz, which would enable the nonseparable terms introduced by the Augmented Lagrangian to be linearized, cannot be implemented (at least in a "large step" version), since the cost function is not differentiable.

The Auxiliary Problem Principle will enable us to find a satisfactory solution by seeking a form of compromise between the two algorithms. Indeed this highly general formalism guarantees the convergence of an algorithmic continuum from gradient methods to the explicit resolutions [6], [5].

More specifically, [7] by linearizing at each iteration the nonseparable terms introduced by the Augmented Lagrangian and a part of the criterion proves the following convergence result.

Proposition 1: Consider the problem (9). Also suppose that:

- $J = J_1 + J_2$, where J_1 is a convex, differentiable function with its derivative Lipschitz of constant A over U^{ad} and where J_2 is a lower semicontinuous (l.s.c.) convex function;
- Θ is a τ-Lipschitz function,
- the Lagrangian $\mathcal{L}_c$ (10), has a saddle-point over $U^{\text{ad}} \times \mathcal{C}^*$ (here, the condition shall be met if J is coercive and if the constraints qualification conditions are met).

Consider the following algorithm:

$$\min_{u \in U^{ad}} \left[\begin{array}{c} \dfrac{K(u)}{\epsilon} - \left\langle \dfrac{K'(u^k)}{\epsilon}, u \right\rangle + \langle J_1'(u^k), u \rangle \\ + J_2(u) + \langle \lambda^k + c\Theta(u^k), \Theta(u) \rangle \end{array} \right] \quad (11)$$

$$\Downarrow$$

$$u^{k+1}$$

$$\lambda^{k+1} = \lambda^k + \rho\Theta(u^{k+1}) \quad (12)$$

where: $K: \mathscr{U} \to \mathbb{R}$ is a b-strongly convex function, differentiable function over U^{ad} with Lipschitz derivatives.

Then the solution of (11) exists and is unique.

If in addition we assume that $0 < \epsilon < (b/A + c\tau^2)$, and $0 < \rho < 2c$, the sequence $((u^k, \lambda^k))_{k \in \mathbb{N}}$ is bounded and has at least a cluster point in the weak topology. Furthermore, any cluster point of this sequence is a saddle-point of $\mathscr{L}_c$ over $U^{ad} \times \mathscr{C}^*$.

This result, if we choose $J = J_1$ and $K: u \mapsto \frac{1}{2}\|u\|^2$, proves the convergence of algorithms of Arrow–Hurwicz type.

Choosing $J_2 = J$, however, gives a form of algorithm which is very close to that of Uzawa. In the latter case, only the terms introduced by the Augmented Lagrangian will be linearized. Of course, the primal minimization step which has to take explicitly into account the function J, will thus be more complex than in the case where the latter is linearized. Nevertheless, the differentiability of the cost function does not have to be assumed (the proposition above only requires that J_2 be l.s.c.).

A.4) Application to Daily Generation Scheduling

In order to apply the convergence result (1), we introduce slack variables e_t and we replace constraints (3) by $R_t - \sum_{i \in \mathscr{I}} r_{i,t} + e_t = 0$ and $e_t \geq 0$ for all $t \in \mathscr{T}$.

From this formulation, by choosing an auxiliary function $K: x \mapsto \epsilon K_\alpha \|x\|^2$ (with $K_\alpha \in \mathbb{R}$) and $J = J_2$, Proposition 1 yields the following algorithm (A^{aug}) (iteration $k + 1$):

$$\min_{p_{i,\mathscr{T}} \in \mathscr{P}_i} \sum_{i \in \mathscr{I}_2} \left[\begin{array}{c} c(p_{i,\mathscr{T}}) - \sum_{t \in \mathscr{T}} (\lambda_t^{k+1/2} p_{i,t} + \mu_t^{k+1/2} r_{i,t}) \\ + K_\alpha \sum_{t \in \mathscr{T}} \left((p_{i,t} - p_{i,t}^k)^2 + (r_{i,t} - r_{i,t}^k)^2 \right) \end{array} \right]$$

$$\Downarrow$$

$$p_{I2,\mathscr{T}}^{k+1}, r_{\mathscr{I}_2,\mathscr{T}}^{k+1};$$

$$\forall t \in \mathscr{T}$$

$$e_t^{k+1} = \max\left\{ \left(e_t^k - \dfrac{\mu_t^{k+1/2}}{2K_\alpha} \right), 0 \right\},$$

$$\begin{cases} \lambda_t^{k+1} = \lambda_t^k + c\left(D_t - \sum_{i \in \mathscr{I}_2} p_{i,t}^{k+1} \right), \\ \mu_t^{k+1} = \mu_t^k + c\left(R_t - \sum_{i \in \mathscr{I}_2} r_{i,t}^{k+1} + e_t^{k+1} \right); \end{cases}$$
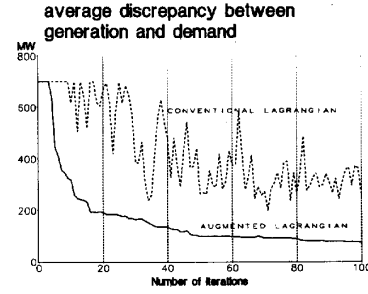

Fig. 5. Evolution of the mean discrepancy between generation and demand.

where

- $\lambda_t^{k+(1/2)} = \lambda_t^k + c(D_t - \sum_{i \in \mathscr{I}_2} p_{i,t}^k)$,
- $\mu_t^{k+(1/2)} = \mu_t^k + c(R_t - \sum_{i \in \mathscr{I}_2} r_{i,t}^k + e_t^k)$,
- $\mathscr{I}_2$ is the set of thermal units and "hydraulic groups."

We have already underlined the output of a hydro-plant is considered as being a piecewise linear function of the discharge. The hydraulic groups here introduced are related to the "linear pieces" of this function. Consequently, the hydro-output may be considered as a linear function of the discharge of these groups.

The introduction of these groups aims at guaranteeing the convexity of the cost function. If we wrote the algorithm (A^{aug}) using the set $\mathscr{I}$, $p_{i,t}$ would be the generation of plant i and thus a piecewise linear function of the discharge. Therefore, nothing would prove the quadratic terms $(p_{i,t} - p_{i,t}^k)^2$ to be convex functions of the discharge.

This algorithm has been compared to that of Uzawa. The comparison has only concerned the set of the thermal plants. All the operating constraints which have been stipulated hereabove were taken into account.

Figure 5 represents the evolution, during the iterations, of the mean deviation $\bar{E}$ between generation and demand: $(1/\text{car } d\mathscr{T})\sum_{t \in \mathscr{T}} |\sum_{i \in \mathscr{I}_2} p_{i,t} - D_t|$.

These curves show that:

- despite the nonconvexity of the local problems, the Augmented Lagrangian steadies the convergence of the algorithm,
  - the speed of convergence is greatly increased,
  - the solution which is reached is closer to satisfying the coupling constraints.

In order to measure the efficiency of the algorithm, it may be observed that the discontinuities in the operating range of nuclear power plants are of the order of 200 MW. Thus the discrepancy between generation and demand which is achieved by this technique is therefore of the order of half these discontinuities: using an algorithm whose convergence proof is based on convexity assumptions, one could hardly hope for greater efficiency.

Furthermore, the inequality constraints corresponding

to the requirements on the power reserve can be perfectly satisfied by this algorithm.

In this way, despite the nonconvexities inherent to the problem that we have dealt with, the numerical tests achieve what the theory promised.

Over and above this result, however, the efficiency of the coordination process implemented here opens up other channels. These are based on reformulating the original problem by "splitting" the variables.

A.5) Splitting Variables

"Splitting variables" consists in substituting for the resolution of a problem of the type:

$$\min_{u \in U^{ad} \cap V^{ad}} J_1(u) + J_2(u), \qquad (13)$$

where:

• J_1: $\mathscr{U} \to \mathbb{R}$ and J_2: $\mathscr{U} \to \mathbb{R}$ are two l.s.c. convex functions,

• U^{ad} and V^{ad} are two closed convex sets of the Hilbert space $\mathscr{U}$,

the problem:

$$\min_{u \in U^{ad}, v \in V^{ad}} J_1(u) + J_2(v), \qquad (14)$$

subject to

$$u - v = 0. \qquad (15)$$

The interest of this transform only appears if we consider processing by dual methods these constraints (15).

The algorithm (11) applied to (14) yields (iteration $k + 1$):

$$\min_{u \in U^{ad}} \left(\frac{K_1(u)}{\epsilon} - \left\langle \frac{K_1'(u^k)}{\epsilon}, u \right\rangle + \left\langle \lambda^{k+1/2}, u \right\rangle + J_1(u) \right)$$

$$\Downarrow$$

$$u^{k+1},$$

$$\min_{v \in V^{ad}} \left(\frac{K_2(v)}{\epsilon} - \left\langle \frac{K_2'(v^k)}{\epsilon}, v \right\rangle - \left\langle \lambda^{k+1/2}, v \right\rangle + J_2(v) \right)$$

$$\Downarrow$$

$$v^{k+1},$$

$$\lambda^{k+1} = \lambda^k + \rho(u^{k+1} - v^{k+1})$$

where

• $\lambda^{k+1/2} \stackrel{\text{def}}{=} \lambda^k + c(u^k - v^k)$,

• K_1 and K_2 are two strongly convex functions (with respective constants b_1 and b_2), differentiable with Lipschitz derivatives.

Proposition 1 proves the convergence of this algorithm providing that $b_1 > 2c, b_2 > 2c$ and $0 < \rho < 2c$.

It may be noted that the constraints defining U^{ad} and V^{ad} do not have to be taken into account jointly here. It is

therefore well and truly a decomposition mode by "constraints breaking" which is achieved by this transform.

A.6) Application to Daily Generation Scheduling

This splitting variables transform has been applied to the daily generation scheduling problem. Its aim here is to avoid the simultaneous processing of the space and time dimensions of the problem being dealt with. Within this framework, problem (1) is transformed as follows:

$$\min \sum_{i \in \mathscr{I}_2} c(p_{i,\mathscr{T}}) \qquad (16)$$

subject to

$$\forall t \in \mathscr{T}: \begin{cases} \displaystyle\sum_{i \in \mathscr{I}_2} q_{i,t} = D_t, \\ \displaystyle\sum_{i \in \mathscr{I}_2} s_{i,t} = R_t \\ \forall i \in \mathscr{I}_2: \begin{cases} q_{i,t} \in \left[\underline{P_i}, \overline{P_i}\right] \\ s_{i,t} \in \left[0, \overline{P_i} - q_{i,t}\right] \end{cases} \end{cases} \qquad (17)$$

$$\forall i \in \mathscr{I}_2 : p_{i,\mathscr{T}} \in \mathscr{P}_i \qquad (18)$$

$$\forall i \in \mathscr{I}_2, \qquad \forall t \in \mathscr{T}: p_{i,t} = q_{i,t}, \qquad r_{i,t} = s_{i,t}. \qquad (19)$$

The dynamic constraints related to the operation of power plants (18) concern the variables p and r, while the variables q and s take into account the spatial component: the coupling constraints (17).

It may be noted that constraints $q_{i,t} \in [\underline{P_i}, \overline{P_i}]$ and $s_{i,t} \in [0, \overline{P_i} - q_{i,t}]$ are not needed to achieve a formulation equivalent to (1). These constraints are necessarily met by the variables p and r, and are therefore met by q and s according to (19). Nevertheless, handling these constraints has a very low cost (in computing time) and has a "structuring" effect.

The number of splitting constraints (19) which have to be processed by duality is quite considerable. There are more than 10000 "splitting" constraints: hundreds of plants and 48 times steps.

Therefore, despite the efficiency of the coordination method described above, the advantage of such a transform may seem quite dubious.

Nonetheless, the algorithm (A^{ded}) can be applied to solve (16):

$$\min_{p_{i,\mathscr{T}} \in \mathscr{P}_i} \sum_{i \in \mathscr{I}_2} \left\{ \begin{array}{l} c(p_{i,\mathscr{T}}) - \displaystyle\sum_{t \in \mathscr{T}} (\lambda_{i,t}^{k+1/2} p_{i,t} + \mu_{i,t}^{k+1/2} r_{i,t}) \\ + K_\alpha \displaystyle\sum_{t \in \mathscr{T}} \left((p_{i,t} - p_{i,t}^k)^2 + (r_{i,t} - r_{i,t}^k)^2 \right) \end{array} \right\}$$

$$\Downarrow$$

$$p_{\mathscr{I}_2,\mathscr{T}}^{k+1}, r_{\mathscr{I}_2,\mathscr{T}}^{k+1}$$

$$(20)$$

$\forall t \in \mathscr{T}$:

$$\min_{q_{i,t}, s_{i,t}} \left[\begin{array}{l} \sum_{i \in \mathscr{I}_2} \left(\lambda_{i,t}^{k+1/2} q_{i,t} + K_\alpha (q_{i,t} - q_{i,t}^k)^2 \right) \\ + \sum_{i \in \mathscr{I}_2} \left(\mu_{i,t}^{k+1/2} s_{i,t} + K_\alpha (s_{i,t} - s_{i,t}^k)^2 \right) \end{array} \right]$$

subject to:

$$q_{i,t} \in \left[\underline{P}_i, \overline{P}_i \right], s_{i,t} \in \left[0, \overline{P}_i - q_{i,t} \right]$$

$$\sum_{i \in \mathscr{I}_2} q_{i,t} = D_t, \sum_{i \in \mathscr{I}_2} s_{i,t} = R_t \qquad (21)$$

$$\Downarrow$$

$$q_{\mathscr{I}_2, t}^{k+1}, s_{\mathscr{I}_2, t}^{k+1},$$

$\forall i \in \mathscr{I}_2, \forall t \in \mathscr{T}$:

$$\lambda_{i,t}^{k+1} = \lambda_{i,t}^k + c(q_{i,t}^{k+1} - p_{i,t}^{k+1}),$$

$$\mu_{i,t}^{k+1} = \mu_{i,t}^k + c(s_{i,t}^{k+1} - r_{i,t}^{k+1}),$$

where for all $(i, t) \in \mathscr{I}_2 \times \mathscr{T}$:

- $\lambda_{i,t}^{k+1/2} = \lambda_{i,t}^k + c(q_{i,t}^k - p_{i,t}^k),$
- $\mu_{i,t}^{k+1/2} = \mu_{i,t}^k + c(s_{i,t}^k - r_{i,t}^k).$

The algorithm turns out to be remarkably efficient. The curves (Fig. 6) which, as above, compare the evolution of the mean discrepancy between generation and demand during the iterations, show that the criterion decreases steadily.

Moreover, the reached solution is, on average, less than 100 MW from the exact satisfaction of the constraints linking generation and consumption. Therefore, despite the considerable number of dualized constraints, the convergence of this algorithm is faster and more accurate than (A^{aug}).

Furthermore, it can be noted that the volume of computations to be performed at each iteration is quite comparable to that necessary for one iteration of (A^{aug}). Indeed, the subproblems dealing with dynamic constraints (20) have exactly the same structure as those appearing in (A^{aug}). The "static" subproblems (21) are very simple quadratic problems: their resolution requires less than one percent of the computing time necessary to solve (20).

The interest of this method is thus clear from a numerical point of view. But this decomposition strategy has another advantage: it enables additional coupling constraints such as transmission constraints to be introduced.

The transmission constraints are in fact "static constraints." In the direct current approximation, they are modelled as inequality constraints on linear combinations of unit outputs. Denoting $\mathscr{D}_t$ the domain thus defined for each time t, the introduction of additional constraints such that $q_{i,t} \in \mathscr{D}_t$ in (17) enables, at least in principle, such constraints to be processed (the algorithm (A^{ded}) may also be considered as corresponding to a single node network).

Nevertheless, there remain two questions:

- Does taking transmission constraints into account perturb convergence?

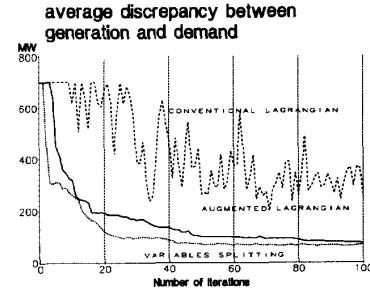average discrepancy between generation and demand

Fig. 6.   Evolution of the mean discrepancy between generation and demand.

- Will the processing of these new constraints, at each iteration at each time step in the static subproblems (17), not lead to an unacceptable deterioration in the performance of the model?

To the first question, after tests have been performed, handling transmission constraints at one step t_0 [2], the answer is negative. The convergence speed and the quality of the solution reached in this case, are quite similar to that which has already been presented: convergence is achieved within some forty iterations, and the generation schedules thus found respect the transmission constraints to within 100 MW.

There remain, however, difficulties linked to performance. A small-scale representation of the network (a few hundred nodes) in fact results in introducing tens of thousands of constraints. The security constraints referred to as "$N - 1$" must be taken into account. These aim at ensuring the consumption could be met even in the event of loss of a transmission circuit. They thus lead to handle all the constraints related to each of the corresponding downgraded situation. From it the astronomical number of constraints to be processed.

Noting that at each time step and at each iteration, a "static" problem has to be solved, clearly solving method must be particularly efficient.

The structure of these static problems, however, is well known to specialists of network problems. It is a problem of "Optimal Power Flow" (OPF) with security constraints, and it has been possible to develop particularly efficient algorithms in this field [4], [8]. Of course the work done so far does not provide answers to all difficulties encountered here (the reader is reminded that some 5000 OPF have to be solved). But the know-how gained in this field will be of considerable help in developing algorithms suited to these repetitive resolutions. Work is already under way in the field and should enable the first life-size experiments to be carried out within the next two years.

We may therefore see here the considerable advantage provided by the joint use of the Augmented Lagrangian technique and the splitting variable transform. This formalism, over and above its numerical qualities and its flexibility enables the full use of know-how gained in each of the fields which we are trying to coordinate.

On the shorter term, however, the APOGEE program, which will be available for use by the end of 1993, will make use of these experiments by integrating a simplified representation of the network.

B. The APOGEE Software

This program, which is currently at a prototype stage, implements the techniques of the Augmented Lagrangian and splitting variables. The formulation (1) of the generation scheduling problem is completed by the introduction of some transmission (static) constraints corresponding to an aggregated network.

The algorithm which is then achieved is the same as (A^{ded}) except for the fact that static subproblems (21) must handle some transmission constraints. Since the number of these coupling constraints remains measured, the resolution of the "static" problems will therefore not pose any technical difficulties.

Compared with those obtained by using a usual Lagrangian technique (5), the subproblems dealing with dynamic constraints see their formulation somewhat modified: the Augmented Lagrangian introduces quadratic terms. As a result, the resolution mode of the hydraulic problems which greatly depends on the linear nature of the problem has to be updated.

B.1) The Hydro Subproblems: Technique of Projected Reduced Gradient

The hydraulic subproblems consist in the minimization of a quadratic function f—the Auxiliary Problem Principle introduces quadratic terms, subject to linear constraints having a flow structure (see II). These optimization problems are solved by a projected reduced gradient technique [9].

Denoting:

- A the $m \times n$ constraint matrix,
- I^k the set of the basic columns at iteration k (and $\overline{I^k}$ the complementary set),
- x^k the current value of the variables,
- $\nabla f(x)$ the gradient of f at point x,
- M^* the transpose of matrix M.

the main steps of this algorithm are the following:

- *computation of the reduced gradient*:

$$\bar{g}^k = \nabla f(x^k)_{\overline{I^k}} - \nabla f(x^k)_{I^k}(A^{I^k})^{-1} A^{\overline{I^k}},$$

- *projection of $\bar{g}^k$*:

$$\forall i \in \overline{I^k}: \pi(\bar{g}^k)_i = \begin{cases} 0 & \text{if } \left(\bar{g}_i^k > 0 \text{ and } x_i^k = \underline{x}_i\right) \\ & \text{or } \left(\bar{g}_i^k < 0 \text{ and } x_i^k = \bar{x}_i\right) \\ \bar{g}_i^k & \text{if not} \end{cases}$$

- *calculation of a descent direction*:
  * if $\theta_{\text{opt}}^{k-1} < \theta_{\text{max}}^{k-1}$ ($\theta_{\text{max}}^{k-1}$ and $\theta_{\text{opt}}^{k-1}$ are defined at the next step of the algorithm), $d^k_{\overline{I^k}}$ is the conjugate direc-

tion:

$$d^k_{\overline{I^k}} = -\pi(\bar{g}^k)_{\overline{I^k}} + \frac{\|\pi(\bar{g}^k)_{\overline{I^k}}\|^2}{\|\pi(\bar{g}^{k-1})_{\overline{I^k}}\|^2} d^{k-1}_{\overline{I^k}};$$

if not $d^k_{\overline{I^k}} = -\pi(\bar{g}^k)_{\overline{I^k}}$;

* $d^k_{I^k} = -(A^{I^k})^{-1} A^{\overline{I^k}} d^k_{\overline{I^k}}$.

- *minimization of the criterion in the direction d^k*:
  * calculation of the maximal moving θ^k_{max} in the direction d^k leading to a feasible solution,
  * $\theta^k_{\text{opt}} = \min(\theta^k_{\text{max}}, -\langle d^k, \nabla f(x^k)\rangle / \langle d^k, \nabla^2 f(x^k) d^k\rangle)$,
  * $x^{k+1} = x^k + \theta^k_{\text{opt}} d^k$;
- *updating the basis if necessary*: if $(\theta^k_{\text{opt}} = \theta^k_{\text{max}})$ and if a variable i which reaches its bound is a basic variable (if $i \in I^k$) then i leaves the basis and a variable j such that $d^k_i (A^{I^k})_i^{-1} A^j d^k_j < 0$ enters the basis.

Moreover it may be noted that, as in the linear case, the only *a priori* complex calculations which have to be done at each iteration are related to the resolution of the systems $A^{I^k} x = b$ and $\lambda A^{I^k} = c$. Therefore, the flow structure of the matrix A is used in the same way as in the implementation of the simplex algorithm (cf. Section II).

B.2) The Thermal Subproblems: Accelerating Dynamic Programming

The computing cost of dynamic programming becomes prohibitive as soon as the dimension of the state vector increases. In order to overcome this drawback, the local thermal problems are solved as follows:

- the daily constraints [C1), C2), C2)'] are processed by duality,
- the number of states necessary to take into account duration constraints (minimum duration of...) is reduced by considering the problem as an optimization problem on a graph.

Given the specific characteristics of the problem and the low number of constraints to process by duality, the number of iterations to perform is often low and generally zero (the power plant having "no interest" in overriding these daily constraints). At each ("large") iteration, a feasible solution is generated. The feasible output corresponding to the greatest value of the dual function is chosen.

To process constraints of the type minimum duration of... (minimum duration of shut-down for example), the conventional approach consists in increasing the dimension of the state vector. A dimension is introduced containing data on the time already taken, information which enables to return to markovian situation: the state becomes an exhaustive summary of the past (or the future).

This process has the inconvenience of considerably increasing the number of states and therefore of having a considerable effect on computing time.

Nevertheless, the Bellman value may be computed as

follows:

$$V_t(x) = \min_{u \in \mathcal{U}^d(x,t)} c^d(x,u,t) + V_{t+d(u,t)}(f^d(x,u,t)),$$

where

- $d(u,t)$ is the duration of the transition related to the control u at time t,
- $f^d(x,u,t)$ is the system state at time $t + d(u,t)$ assuming that control u is applied to the system which at time t is in the state x,
- $c^d(x,u,t)$ is the operating cost of the system during the transition from the state x at time t to the state $f(x,u,t)$ at time $t + d(u,t)$.

To deal with the minimum down time constraints d^{dur}, for example, using the above-mentioned equation will lead to multiply by two the number of states. On the other hand, to handle this constraint in a conventional way would multiply by $d^{dur} + 1$ the number of states. This trick thus allows to save a considerable part of the storage and of the computing time.

IV. EXPECTED IMPROVEMENTS: TOWARDS REAL TIME OPTIMIZATION

The lapse of one day between two updatings of daily generation schedules may seem quite long. As we have already underlined, it basically corresponds to the data processing capacity available today. Implementing such a function during the day in fact requires full information on the status of the power plants and a software and hardware infrastructure enabling the fast and reliable (computerized) transmissions of controls to each of the plants. The system which has to be set up is therefore complex and relatively expensive. Given the implications (economic but also in terms of operating security), the decision to develop such a system has been taken by Electricité de France: the decision has given rise to the CASOAR program which includes a research and development program aiming at designing new families of control centers implementing real-time optimization.

As part of this program, forward-looking functions for generation operation will be progressively implemented into the national control center. In this way, the optimization of generation schedules could be performed several times a day.

Is there, however, a real interest in performing this type of review on-line? While it is difficult to provide any quantitative evaluation of savings we might expect, the increasingly stochastic nature of the problem allows us to surmise that these savings would be quite substantial, over a time scale of several years.

A noteworthy increase in the uncertainty of demand forecasts can be observed in France. This corresponds to an evolution of the consumption modes: during winter time the proportion of electric heating in consumption is greater every year. Furthermore, as we have already emphasized, temperature forecasts to within one day integrate a certain factor of error (a typical deviation of some 2°C will lead to a deviation of 2500 MW).

Therein lies the interest of optimizing generation on-line.

A development of this nature does, however, pose certain difficulties.

Drawing up generation schedules the day before allows constraints to be neglected by the corresponding optimization models. In real time, however, the proposed procedure must systematically give reliable, feasible and economically satisfactory results.

To achieve such an accuracy, the algorithm developments presented above are a considerable advantage since they take into account a large number of additional constraints (such as transmission constraints) and thus achieving a modelization much closer to "reality."

Moreover, the use of mathematically sound methods will provide actual guarantees of reliability to the proposed procedure. From experience empirical methods only meet situations that have been expected *a priori*. The scale of the problem and the number of such situations in this case make any such indexing impossible.

The work on algorithms which has been presented in this paper is continuing, in order to define a computer tool which can be used on-line. What we are seeking today to define is:

- an efficient method for repeated resolutions of network (static) subproblems (OPF),
- a process enabling us to take into account all the constraints (and in particular hydraulic constraints).

While the methodological developments are of great interest, it would be illusory to believe that they provide a global answer to the difficulties created by upgrading towards real time.

When decisions are taken on the basis of studies, the end-user does not limit himself to providing the useful data. As a rule, he integrates on the one hand the missing factors in the representation provided by the model, and on the other, translates the results obtained so that they can be used (by other optimization systems or functions which may have another "vision" of the world).

It is thus the end-user who has to manage the inadequacy (even if this is only relative) between modeling and physical reality, as well as the breaks between the representation modes used in each part of the system.

On-line processing does not enable such "arrangements." It is a *continuum of explicit representations* which has to be developed in order to enable the definition of an integrated system.

Given the complexity of a process control system, in which measurement operation, automaton control, network security operation etc., all have to work together, the operation is far from simple.

If we try to pin down the exact nature of the difficulties encountered, we can situate them on two distinct levels.

Firstly, we have to obtain data processing compatibility which is as wide-reaching as possible. The same data do not have the same "worth" depending on the time scale in question (in relation to the decisions that have to be

taken). Consequently their coding could be quite different. Homogeneization is therefore a prerequisite to defining such a system.

The operation is, in fact, particularly difficult since it is the moment when are made explicit not only the undefined but on the same occasion the numerous *a priori* assumptions which are often incompatible.

In addition, the fact that representations of the system are different is not only the result of an "accident" or of negligence. At present, the (on-line) decision which are taken in order to modify a generation schedule depends on the operator judgement. Moreover, the generation mix is not actually automated. Therefore, the operation rules are defined in this context: they aim at ensuring a safe "human" operation.

Just as the functioning of cars does not define the highway code, these operation rules do not directly correspond to the operating constraints of the plants.

Every "daily" constraints we described above are not "actual" constraints: how can a nuclear power plant "know" when it is 24 o'clock?

Parallel to it, a complete change of the operating mode has to be excluded. A straight automation of the whole system is not possible. The necessary upgrading of the software and hardware systems will be progressively implemented.

We thus have to face *a priori* contradictory objectives: to integrate optimization functions without deeply modifying the modelings which have been designed for human operation.

In order to match these objectives, softwares allowing to build up contexts fitted to each optimization function have to be designed. The initialization of an APOGEE type software will require for example the rebuilding of an initial state (in the sense of the dynamic programming). This state is not an intrinsic notion, daily constraints not being handled by very short term models. It will thus be a coherent set of identification and optimization functions that will have to be designed to establish generation plans in such a context. Work is currently under way on this purpose.

V. CONCLUSION

New algorithms we have sketched here represent an important step towards a better economic performance. Nevertheless, they will not be sufficient to ensure an evolution towards real time. In order to make an efficient use of such methods and softwares a modernization of the whole system is required.

First of all upgrading the data processing (measurement and transmission system, data bases etc . . . ,) is a necessary condition. Sophisticated optimization function would be of no use without reliable information about generation units and transmission lines.

Furthermore, the work of the system operators will necessarily change. As a consequence, they will have to be trained in this aim.

This modernization will at every level of the firm is the

necessary condition of a success in this field. And lack of this will is in our mind the main explanation of the failure of many attempts to use mathematically sound methods practically.

Electricité de France has chosen this difficult path of modernization: the decision has given rise to the CASOAR program. In such a context, if the idea, quite simplistic but usual, that an algorithm can completely solve a practical problem is actually false, the methodological developments we have presented hereabove are a considerable advantage.

The Augmented Lagrangian and Splitting Variables techniques enable to notably enrich the modeling of the daily generation scheduling problem we are dealing with: transmission together with operating constraints will be coped with. Moreover, the theoretical background of our approach will guarantee the reliability of the achieved result; guarantee that no heuristic method can provide.

REFERENCES

[1] K. Arrow, L. Hurwicz, and H. Uzawa, *Studies in Linear and Nonlinear Programming.* Stanford: Stanford University Press.

[2] J. Batut and A. Renaud, "Daily Generation Scheduling with Transmission Constraints: A New Class of Algorithms," *IEEE Trans. Power Syst.* vol. 7, no. 3, pp. 982–989, Aug. 1992.

[3] J. Batut, A. Renaud, and P. Sandrin, "A new software for generation rescheduling in the future EDF national control center," in *Proc. PSCC Graz,* 1990.

[4] J. Carpentier, "Optimal power flows," (survey) *Int. J. Elect. Power Energy Syst.* vol. 1, pp. 8–15, Apr. 1979.

[5] G. Cohen, "Auxiliary problem principle and decomposition of optimization problems," *J. Optimiz. Theory Appl.* vol. 32, no. 3, Nov. 1980.

[6] G. Cohen, "Optimization by decomposition and coordination: An unified approach," *IEEE Trans. Automat. Contr.* vol. AC-32, no. 2, Apr. 1978.

[7] G. Cohen and D. L. Zhu, "Decomposition-coordination methods in large scale optimization problems. The nondifferentiable case and the use of augmented Lagrangians," *Adv. Large Scale Syst. Theory and Application,* vol. 1, CRUZ J. B., Ed. Greenwich, CT: JAI Press, Inc., 1983.

[8] H. W. Dommel and W. F. Tinney, "Optimal power flow solutions," *IEEE Trans. Power Apparatus Syst.* vol. PAS-87, Oct. 1968.

[9] P. Faure and P. Huard, "Résolution des Programmes Mathématiques à Fonction non Linéaire par la Méthode du Gradient Réduit," *Revue de Recherche Opérationnelle,* no. 36, pp. 167–206, 1965.

[10] J. F. Maurras, "Optimization of the flow through networks with gains," *Math. Programming,* vol. 3, no. 2, pp. 135–144, 1972.

[11] A. Merlin and P. Sandrin, "A new method for unit commitment at Electricité de France," *IEEE Trans. Power Apparatus Syst.* vol. PAS-102, no. 5, pp. 1218–1225, May 1983.

[12] R. Tyrrell Rockafellar, "A dual approach to solving nonlinear programming problems by unconstrained optimization," *Math. Programming,* vol. 5, pp. 354–373, 1973.
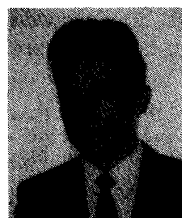
Arnaud Renaud was born in Paris, France, in 1960. He received the Engineer's degree from the "Ecole Nationale Supérieure des Techniques Avancées" in 1983, and graduated in Statistics from Paris VI University, in 1984.

After graduating from Paris VI University, he began working at the Department of Network Planning at the Research and Development Directorate of Electricité de France. In 1988, he joined the Optimization and Simulation Methods Department where he is now interested in the Daily Generation Scheduling problems.