Convergence Properties of the Waveform Relaxation Method as Applied to Electric Power Systems

M. L. Crow, M. D. Ilić, J. K. White

Massachusetts Institute of Technology
Cambridge, MA

Abstract

In this paper several theoretical results pertaining to the application of waveform relaxation to the simulation of power systems are presented.

1 Introduction

The interest in using power system simulation for real-time applications has sparked much research in techniques for fast and accurate simulation of the effects of faults in large power systems[1] [2]. The simulation of power systems usually involves numerically solving large stiff systems of ordinary differential equations (ODEs) subject to algebraic constraints [1]. A recently developed algorithm for such problems, waveform relaxation(WR), shows promise as an efficient technique for power system simulation. In this approach, the ODE and algebraic system is broken into subsystems which are solved independently, each subsystem using guesses about the behavior of the state variables in other subsystems. Waveforms are then exchanged between the subsystems, and the subsystems are resolved with, hopefully improved, information about the other subsystems. This process is continued until convergence is achieved.

Although WR has not proved to be as effective for circuit simulation, for which it was originally developed, as hoped, the algorithm may perform better on power system simulation problems that, under certain assumptions only, can be modeled as ODE's in normal form. Therefore, the commonly experienced difficulties in applying WR to the VLSI circuitry, due in part to floating capacitance, are not anticipated. In this paper we investigate some of the theoretical issues, and examine some simple examples, to try to characterize the behavior of WR when applied to power systems simulation problems that can be cast into ODE's in normal form. In the next section we start by briefly describing WR, in Section 3 we present a generalization of some convergence results for diagonally dominant systems, in Section 4 we examine a model for a slightly unstable power system, and in Section 5, we investigate the numerical stability of multirate integration. Lastly, conclusions and acknowledgment are given in Section 6.

2 The WR Algorithm

The transient behavior of a large coupled system of generators and loads can be reasonably well described by a large coupled system of ordinary differential equations $\dot{x} = f(x,t) \quad x(0) = x_0$ where $x_i \in \Re^n$ is the state corresponding to the power system dynamic model pair $[\delta_i \ \omega_i]$ which are the rotor angle and angular speed of the i^{th} generator respectively. The power system dynamic model is

$$\dot{\delta}_i = \omega_i - \omega_s \tag{1}$$

$$\dot{\omega}_i = \frac{1}{M_i}(P_{m_i} - E_i^2 g_{ii} - E_i \sum_{j \neq i}^{n} E_j (b_{ij} sin\delta_{ij} + g_{ij} cos\delta_{ij})) \tag{2}$$

for $i = 1, \ldots, n$, where $\delta_{ij} = \delta_i - \delta_j, \omega_s = 377$radians per second, M_i is the inertia constant, P_{m_i} is the mechanical input per unit (constant), $E_i \angle \delta_i$ is the constant voltage behind transient reactance, and b_{ij}, g_{ij} are the elements of the reduced admittance matrix. The loads are represented by passive impedances. In particular, this "classical" model gives a good approximation for the behavior of the system during the "first swing", which is of the order of one second following a fault in a typical power system.

To solve the ODE system with the WR algorithm, the system is usually first partitioned into subsystems, where the subsystems are are selected so that tightly coupled state variables are in one subsystem. In particular, decompose the system into m subsystems as

$$\dot{x}_1 = f_1(x_1, x_2, ..., x_m, t) \quad x_1(0) = x_{1_0} \tag{3}$$

$$\vdots$$

$$\dot{x}_m = f_m(x_1, x_2, ..., x_m, t) \quad x_m(0) = x_{m_0} \tag{4}$$

where $x_i \in \Re^{n_i}$, $\sum_{i=1}^{m} n_i = n$, and $f_i : \Re^n \rightarrow \Re^{n_i}$. To shorten the notation, let $f_i(x^{i,k}, x^{i,k+1}, t) \equiv f_i(x_1^k, ..., x_{i-1}^k, x_i^{k+1}, x_i^k + 1..., x_m^k, t)$. This notation is used below to describe the the Gauss-Jacobi WR algorithm applied to solving (4).

Algorithm 1 - Gauss-Jacobi WR Algorithm.
$k \leftarrow 0$.
Guess some $x_i^0(t)$ on $[0,T]$ such that $x_i^k(0) = x_i(0)$.
repeat {
$k \leftarrow k + 1$.
for each $(i \in \{1, .., m\})$ solve on $[0,T]$

$$\dot{x}_i^{k+1} = f_i(x^{i,k}, x^{i,k+1}, t) \quad x_i^{k+1}(0) = x_i(0) \tag{5}$$

} **until** $(\|x^k - x^{k-1}\| \leq \epsilon)$

3 Diagonally Dominant Systems

It has been shown [3] that the WR algorithm applied to (4) will converge over any finite interval to the solution of the

ISCAS '89

differential system for any initial guess which is consistent with the initial conditions, and this convergence has been shown to be superlinear[4]. These general convergence results are based on the same mechanism as the Picard iteration convergence, and predict a very nonuniform type of convergence. That is, each iteration of the relaxation process may only correct the solution in a short interval, where the size of the interval is inversely related to the Lipschitz constant of f, l. Clearly for stiff systems, where by definition the time interval is large compared to the inverse of l, this bodes ill for the efficiency of WR.

If the system exhibits its stiffness on the diagonal, in a sense we will be specific about shortly, it can be shown that the WR algorithm is a contraction in the l_∞ norm on the space of functions $[0,T] \rightarrow \Re^n$. Several results of this type have appeared previously [7], but a stronger result is presented below which applies to nonlinear block-partitioned problems, and is based on very simple proof given in the appendix.

Theorem 1 *If, for the iteration update equations in (5), the inner product*

$$\langle (x_i^1 - y_i^1), (f_i(x^{i,0}, x^{i,1}, t) - f_i(y^{i,0}, y^{i,1}, t)) \rangle < 0 \qquad (6)$$

when $\|(x_i^1 - y_i^1)\|_2 > (1-\epsilon)\|x_{noi} - y_{noi}\|_2$ *where* $x_{noi} \equiv (x_1, ..., x_{i-1}, x_{i+1}, ..., x_m)^T$ *and* ϵ *is a positive number less than 1, then (5) represents a contraction in a uniform norm. That is, for any* $y^0(t)$ *and* $z^0(t)$ *which satisfy the initial conditions of (4)*

$$max_{i\in\{1,...,m\}} max_t \|x_i^1 - y_i^1\|_2 \leq (1-\epsilon) max_{i\in\{1,...,m\}} max_t \|x_i^1 - y_i^1\|_2. \qquad (7)$$

In addition, if (6) holds for $\epsilon \geq 0$, *and in addition if for any* x, y *there exists some* K *such that*

$$\langle (x_i^1 - y_i^1), (f_i(x^{i,0}, x^{i,1}, t) - f_i(y^{i,0}, y^{i,1}, t)) \rangle < K[\|(x_i^1 - y_i^1), (x_i^1 - y_i^1)\|_2^2 - \|(x_i^0 - y_i^0), (x_i^0 - y_i^0)\|_2^2] \qquad (8)$$

WR is still a contraction on any finite interval $[0,T]$,

$$max_{i\in\{1,...,m\}} max_{t\in[0,T]} \|x_i^1 - y_i^1\|_2 \leq \gamma(T) max_{i\in\{1,...,m\}} max_{t\in[0,T]} \|x_i^1 - y_i^1\|_2 \qquad (9)$$

where $\gamma(T) < 1$.

It is easy to see where the condition in (6) stems from by viewing one equation out of the partitioned system in (4),

$$\dot{x}_i = f_i(x_1, x_2, ..., x_m, t) \quad x_i(0) = x_{i_0}, \qquad (10)$$

as an independent differential equation with $m-1$ inputs. The condition in (6) says simply that with respect to x_i, f_i is so dissipative that $max_t \|x_i(t)\|_2$ can never come closer than $1-\epsilon$ of $\sum_{j=1,j\neq i}^n max_t \|x_j\|_2$. Or, loosely, the gain from the sum of the inputs, x_j, $j \neq i$, to the output, x_i, is less than $1-\epsilon$. Finally, in the case where $\epsilon = 0$ the gain from $\sum_{j=1,j\neq i}^n max_t \|x_j\|_2$ to $\|x_i\|_2$ is still strictly less one on any finite interval if (8) is satisfied. Although odd-looking, the condition in (8) just insures that the upper bound is approached like a decaying exponential.

For a system $\dot{x} = f(x,t)$ to satisfy the conditions of Theorem 1 depends both on the characteristics of f, and how it is partitioned. For example, if for all x, the Jacobian of f, $J_f(x) = \frac{\partial f(x)}{\partial x}$, is diagonally dominant with negative diagonal entries, and the system is partitioned into scalar equations, (6) and (8) will be satisfied. In addition, $\epsilon > 0$ if $J_f(x)$ is strictly diagonally dominant uniformly in x. The condition in (6) will not necessarily be satisfied if this same system is partitioned into blocks. In the diagonally dominant case just described, this is an an artifact of the conditions of the theorem. The results of Theorem 1 are still true regardless of the partitioning in this diagonally dominant case, but the authors could only find a very specific and slovenly proof which we omitted so as not to try the reader's patience.

4 Investigation of Unstable Systems

That WR is a contraction in a uniform norm for problems that are diagonally dominant suggests that slight perturbations from this case might still converge reasonably uniformly. This is not necessarily true as can be demonstrated by example. Consider the following 2x2 example that is intended to be a rough approximation to the structure of an unstable power system

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} -1+\epsilon & 1 \\ \alpha & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \qquad (11)$$

Note that this problem is nearly diagonally dominant with negative diagonal elements. When $\epsilon = 0$, system (11) can be thought of as representing equivalent power system dynamics in its stable mode, provided $\alpha < 1$. Parameter $\epsilon \neq 0$ is intended to reflect dynamic instabilities which could be caused by either low voltages or short circuits.

The WR algorithm was used to compute the solution to (11) on $[0,T]$ with initial condition $x(0) = [1\ 0]^T$, $\epsilon = 0.03$, and initial guess $x_i^0(t) = x_i(0), i = 1,2, t \in [0,100]$. This initial guess for the solution contains errors in the fast mode of the system, and as can be seen from the graphs of the WR iterates for x_1 in Figure 1, the iterates converge very nonuniformly to the unstable solution. In Figure 2, the initial condition is $x(0) = [1\ 1]^T$ and the iterates converge in a more uniform manner because the errors in the fast mode have been "killed off" initially. This is a key aspect of WR, that in a practical sense the interval over which there is some uniformity in the convergence is partly a function of the errors in the initial guess, and *can not* be entirely predicted by examining system properties.

Given a linear time-varying system of the form

$$\dot{x}_i = -d_{ii}(t)x_i(t) + \sum_{j\neq i}^n n_{ij}(t)x_j(t) + v_i(t) \qquad (12)$$

where $i = 1, \ldots, n$ and $d_{ii}(t) > 0\ \forall t \in [0,T]$, it is possible to find an estimate a t^*, the size of the interval over which the WR algorithm will be a contraction in the l_∞ norm.

Theorem 2 *If* $d_{ii}(t) < 0$, *then (12) is a contraction in a uniform norm on the finite interval* $[0,t^*]$ *where* $t^* \equiv min_i\{t_i^*\}$ *and*

$$t_i^* = \frac{1}{|\tilde{d}_{ii}|} \left| ln\left(1 - \frac{\gamma|\tilde{d}_{ii}|}{\sum_{j\neq i}^n max_{[0,T]}|n_{ij}(t)|}\right)\right| \qquad (13)$$

where $\tilde{d}_{ii} = min_{[0,T]} d_{ii}(t)$ *and* γ *is the contraction factor.*

Our example indicates that the t^* estimate in Theorem 2 is accurate in predicting the performance in the first situation, predicting a small window during which time the effect

of the stiffness would die out, and the remainder of the simulation could choose larger windows over which the WR would converge in a more uniform manner. In the second case, however, the t^* estimate would be overly conservative. Thus, this t^* estimate is a lower bound on the window size for uniform convergence.

5 Multi-rate Instability

A major advantage of the WR algorithm is that as the differential equations are solved in a decomposed fashion. This implies that if discretization methods are used to solve the independent differential equations the time steps used by the subsystems can be selected fairly independently. Applying this philosophy leads naturally to two questions: does the relaxation still converge, and if it does converge, does the resulting multirate integration method inherit the stability properties of the integration method used for the decomposed systems. The convergence properties of discretized WR algorithms have been discussed elsewhere [5], and in this section we will try to connect WR convergence and multirate stability

Even a carefully implemented backward-Euler based multirate integration method does not necessarily inherit the A-stable properties of backward-Euler. As an example, consider the multirate backward-Euler algorithm with linear interpolation applied to a 2x2 system $\dot{x} = Ax$. The discretized equations are

$$x_i(t+mh) = x_1(t) + mh(a_{11}x_1(t+mh) + a_{12}x_2(t+mh))$$

$$x_2(t+kh) = \frac{ha_{21}}{m}\sum_{i=1}^{k}\frac{i}{(1-ha_{22})^{k-i+1}}x_1(t+mh)$$

$$+\sum_{i=1}^{k}\frac{1-\frac{i}{m}}{(1-ha_{22})^{k-i+1}}x_1(t) + \frac{1}{(1-ha_{22})^{k-i+1}}x_2(t) \quad (14)$$

where $1 \leq k \leq m$, which reduces to the usual backward-Euler algorithm if $m = 1$. Note that this is precisely the multi-rate integration algorithm that would be produced by a convergent discretized WR process. To analyze the region of absolute stability for the multirate method, Gaussian elimination is used to reduce the above equations to the form $\mathbf{x}(t+mh) = M(m,h)\mathbf{x}(t)$.

The region of absolute stability for the multirate integration method is then those values of m and h for which the eigenvalues of $M(m,h) < 1$. If $m = 1$, which is regular backward-Euler, then if the eigenvalues of A have negative real parts, $M(1,h) < 1$. However, that is not true for all m. In particular, if

$$A = \begin{bmatrix} -1 & 1 \\ -0.54 & 0.5 \end{bmatrix} \quad (15)$$

The eigenvalues of A are $-0.1, -0.4$, and $M(100,1)$ has eigenvalues of -0.23 and 2.01, thus the multirate algorithm is unstable.

One justification for the multirate instability can be seen by considering solving the 2×2 system above with the A as in (15) using the WR algorithm. In that case the iteration equations become $\dot{x}_1^k = -x_1^k + x_2^{k-1}$ and $\dot{x}_2^k = 0.5x_2^k - 0.54x_1^{k-1}$. Note that the subsystem of x_2 is unstable with its "eigenvalue" being greater than zero, even though the system as a whole is stable. The effect on the WR iterates will be that unless the initial guess is almost the exact solution, the iterates will never converge with any uniformity over long intervals. Contrast this with the example in Section 4, which was unstable but in which the subsystems where stable, and in certain circumstances, the WR algorithm converged uniformly. Power systems often exhibit this latter characteristic. Typically, a group of generators, which act as a unit going unstable with respect to the entire system, will remain stable with respect to the other generators within the unit. In this scenario, one would expect the WR algorithm to be an efficient method of simulation.

We finish this section with theorem that makes a weak connection between the stability of a multirate integration scheme, and the convergence of WR. In particular it suggests that when partitioning a system it is very important to insure the subsystems are stable.

Theorem 3 *If a stable linear system of the form of (4) is such that when the WR algorithm is applied to solving the system, the iterates contract in a uniform norm, and if the decomposed subsystems are stable and are integrated with an A-stable method, and finally, if the discrete sequences produced by numerically integrating the WR iteration equations converge, the computed solution is within a timestep independent constant times $\frac{1}{1-\gamma}I_{err}$ where γ is the contraction factor for the relaxation and I_{err} is the maximum local truncation error.*

Theorem 3 follows directly from the fact that when integrating a stable linear problem, the global error is a timestep independent constant times the maximum local truncation error, and a little reorganization of the triangle inequality.

6 Conclusions

In this paper several theoretical results were presented, and simple examples examined in order to determine the suitability of WR for transient power system simulation. This examination has led to two practical suggestions that can easily be satisfied for power systems. The first is to break the simulation interval into sections, the first of which should be narrow and be used to kill off errors in the initial guess that activate the stiff modes. The second is to partition the system into subsystems that are stable, as this not only aids convergence but insures multirate numerical stability.

The authors would like to thank Resve Saleh, Peter Sauer, M. A. Pai, and John Wyatt for several very valuable discussions.

References

[1] A. Bose, K. A. Clements, "Real-Time Modeling of Power Networks," *Proceedings of the IEEE*, December 1987.

[2] M. Ilić-Spong, M. L. Crow, M. A. Pai, "Transient Stability Simulation by Waveform Relaxation Methods", *IEEE Transactions of Power Systems*, vol. PWRS-2, no. 4, November 1987.

[3] E. Lelarasmee, A. Ruehli, A. Sangiovanni-Vincentelli, "The Waveform Relaxation Method for the Time Domain Analysis of Large Scale Integrated Circuits", *IEEE Transactions on CAD*, vol. 1, no. 3, July 1982.

[4] U. Miekkala, O. Nevanlinna, "Convergence of Dynamic Iteration Methods for Initial Value Problems", Report MAT-A230, Institute of Mathematics, Helsinki University of Technology, Finland, 1985.

[5] J. K. White, A. Sangiovanni-Vicentelli, *Relaxation Techniques for the Simulation of VLSI Circuits*, Kluwer Academic Publishers, Boston, 1987.

[6] D. R. Wells, "Multirate Linear Multistep Methods for the Solution of Systems of Ordinary Differential Equations", Report No. UIUCDCS-R-82-1093, Department of Computer Science, University of Illinois at Urbana-Champaign, Urbana, IL, July 1982

[7] D. Mitra, "Chaotic, Asynchronous Relaxations for the Numerical Solution of Differential Equations by Parallel Processors", A.T. & T. Bell Laboratories Memorandum, October, 1984.

7 Appendix - Proof of Theorems

The lemma and its corollary below contain the key result for the proof.

Lemma 1 *Given*

$$\dot{x}(t) = f(x(t), y(t), t) \quad x(0) = 0, \tag{16}$$

where $x(t) \in \Re^n$, $y(t) \in \Re^m$ *and* $f : \Re^n \times \Re^m \times \Re \rightarrow \Re^n$ *is uniformly Lipschitz continuous with respect to* x, y *and piecewise continuous with respect to* t. *If* f *is such that the inner product* $max_t\langle z, f(z,w,t)\rangle < 0$ *whenever* $z \in \Re^n$, $w \in \Re^m$ *are such that* $\|z\|_2 > (1-\epsilon)\|w\|_2$ *for some positive* $0 \leq < \epsilon < 1$, *then* $max_t\|x(t)\|_2 \leq (1-\epsilon)max_t\|y(t)\|_2$.

Proof of Lemma 1.

Multiplying both sides of (16) yields $x(t)^T\dot{x}(t) = x(t)^T f(x(t), y(t), t)$, or

$$\frac{d}{dt}\|x(t)\|_2^2 = x(t)^T f(x(t), y(t), t) \tag{17}$$

from which it follows that $\|x(t)\|_2$ is continuous and differentiable. Suppose $\|x(t)\|_2^2 = M > max_t(1-\epsilon)^2\|y(t)\|_2^2$. Let $\tilde{t} = min_t\{\|x(t)\|_2^2 = M\}$. By continuity of $\|x(t)\|_2^2$ there must exist a finite interval $[\tilde{t}-\Delta, \tilde{t}]$ over which $\|x(t)\|_2^2$ is monotonically increasing. It follows that $\dot{x}(\tilde{t}) = lim_{\delta+\rightarrow 0}\frac{\|x(\tilde{t})\|_2^2 - \|x(\tilde{t}-\delta)\|_2^2}{\delta} \geq 0$ which contradicts (17).

Corollary 1 *If the assumptions of Lemma 1 hold, and for each* $z \in \Re^n, w \in \Re^m$ *such that* $\|z\|_2 \leq \|w\|_2$, *there is exists a positive constant* K *such that*

$$max_t\langle z, f(z,w,t)\rangle < K(\|w\|_2^2 - \|z\|_2^2) \tag{18}$$

then for any finite interval $max_{t\in[0,T]}\|x(t)\|_2$ *is less than or equal to* $\gamma(T)max_{t\in[0,T]}\|y(t)\|_2$ *where* $\gamma(T) < 1$.

Proof of Corollary 1.

Corollary 1 follows directly from Lemma 1 if $\epsilon > 0$. If $\epsilon = 0$, from (17) and the assumptions of the corollary

$$\frac{d}{dt}\|x(t)\|_2^2 = \tilde{K}(\|y(t)\|_2^2 - \|x(t)\|_2^2) \tag{19}$$

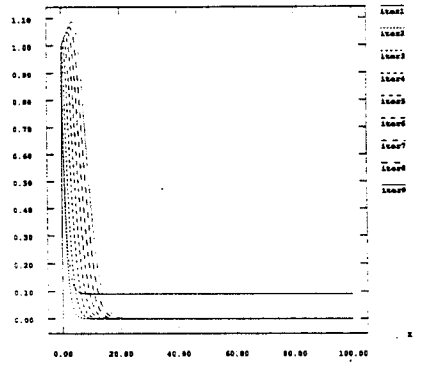

Figure 1: Nonuniformly Converging WR Iterates.

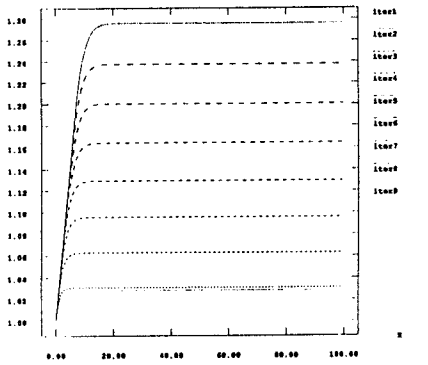

Figure 2: More Uniformly Converging WR Iterates.

where $\tilde{K}$ is the maximum value K in (18), which exists as $x(t)$ is bounded by Lemma 1. Therefore $\|x(t)\|_2^2 \leq (1 - e^{-\tilde{K}t})max_{t\in[0,T]}\|y(\tau)\|_2^2$ which proves the corollary.

Proof of Theorem 1

To prove Theorem 1, from (5), the difference between WR iterations applied to different initial guesses yields

$$\dot{x}_i^1 - \dot{y}_i^1 = f_i(x^{i,0}, x^{i,1}, t) - f_i(y^{i,0}, y^{i,1}, t) \quad x^1(0) - y^1(0) = 0. \tag{20}$$

If f_i is such that the inner product $\langle (x_i^1 - y_i^1), f_i(x^{i,0}, x^{i,1}, t) - f_i(y^{i,1}, t)\rangle < 0$ when $\|(x_i^1 - y_i^1)\| > (1-\epsilon)\|x_{noi} - y_{noi}\|$ where $x_{noi} = (x_1, ..., x_{i-1}, x_{i+1}, ..., x_m)^T$ for any x^0, x^1, y^0, y^1 then (20) satisfies the assumptions of Lemma 1, and the result follows directly.

Proof of Theorem 2

Introducing WR, (12) becomes

$$\dot{x}_i^{k+1}(t) = -d_{ii}(t)x_i^{k+1}(t) + \sum_{j\neq i}^{n} n_{ij}(t)x_j^k(t) + v_i(t) \tag{21}$$

By taking the difference between the iteration $k+1$ and k and reassembling into matrix form, $\|\delta x^{k+1}(t)\| \equiv \|x^{k+1}(t) - x^k(t)\|$ is bounded by

$$max_i(1 - e^{-\tilde{d_{ii}}t})\frac{1}{|\tilde{d_{ii}}|}\sum_{j\neq i}^{n} max_{[0,T]}|n_{ij}(t)|\|\delta x^k(t)\|. \tag{22}$$

The equation (22) will be a contraction in a uniform norm if the interval is less than the t^* given in (13), thus proving the theorem.